REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
06-07-2015	Final Report	1-Apr-2012 - 31-Dec-2012

4. TITLE AND SUBTITLE
Final Report: Dynamic multitasking countermeasures to improve sustained attention while driving

5a. CONTRACT NUMBER
W911NF-12-1-0139

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Paul Atchley, Pascal Doebeck

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Kansas
2385 Irving Hill Road

Lawrence, KS 66044 -7552

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
60733-LS-II.2

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The unknown at the outset of this work was if high cognitive load environments would decrease alertness faster or more slowly. One outcome of the work is that stimulating environments seem to help preserve alertness. However, engaging someone with additional tasks to increase their alertness also resulted in losses to attentional orienting, meaning they missed more information from their environment as the demands of the secondary task increased. Thus, it appears that using secondary tasks to “energize” soldiers may provide benefits to alertness if they are engaged in monotonous tasks, but these tasks may decrease attention to the environment. The conclusion is that in

15. SUBJECT TERMS
vigilance, multitasking

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Paul Atchley
					19b. TELEPHONE NUMBER
					785-864-9803

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Final Report: Dynamic multitasking countermeasures to improve sustained attention while driving

ABSTRACT

The unknown at the outset of this work was if high cognitive load environments would decrease alertness faster or more slowly. One outcome of the work is that stimulating environments seem to help preserve alertness. However, engaging someone with additional tasks to increase their alertness also resulted in losses to attentional orienting, meaning they missed more information from their environment as the demands of the secondary task increased. Thus, it appears that using secondary tasks to “energize” soldiers may provide benefits to alertness if they are engaged in monotonous tasks, but these tasks may decrease attention to the environment. The conclusion is that in cases of the most extreme declines in alertness, a secondary task is beneficial, but generally secondary tasks pose a threat to overall attentiveness.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received		Paper
07/06/2015	1.00	Atchley, Paul, Deboeck, Pasccal, Chan, Mark,, Geldhof, John, Fries, Chelsie. Using momentary derivative estimates to gauge driver performance, Advances in Transportation Studies. , . : ,

TOTAL: **1**

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received	Paper

TOTAL:

Number of Manuscripts:

Books

Received Book

TOTAL:

Received Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Fries, Chelsie	0.25	
Priester, Geoff	0.25	
Chan, Mark	0.25	
FTE Equivalent:	**0.75**	
Total Number:	**3**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Atchley, Paul	0.12	
Deboeck, Pascal	0.12	
FTE Equivalent:	**0.24**	
Total Number:	**2**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: 0.00

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME	
Chan, Mark	
Total Number:	**1**

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

See report

Technology Transfer

Final Report

Grant number: W911NF-12-1-0139

Project title: Dynamic multitasking countermeasures to improve sustained attention

Project investigators: Paul Atchley, Ph.D. & Pascal Deboeck, Ph.D.

Project goals

1) *To better understand potential risks and benefits of multitasking on sustained attention.* The goal of this proposal is to begin to map out the profile of risks and potential benefits for multitasking on sustained attention.

2) *Build techniques based on dynamic systems modeling that are more sensitive than traditional approaches to analyze performance and changes in performance accompanying declines in sustained attention.* Traditional analyses of sustained attention use statistical techniques that are not sensitive to moment-by-moment changes in performance. We predict effects are relatively localized in time and effects are only clear when considering changes across time.

3) *Using the dynamic models we will develop a preliminary understanding how sustained attention differs across individuals as well as within an individual over time; analyses will focus on whether training and prior experience can improve sustained performance and the extent in time that these differences manifest themselves.* We will use dynamic modeling to identify the characteristics of drivers able to successfully sustain attention as a bridge to future work on training to maintain attention under demanding conditions.

Report completed 12 July 2013

Table of Contents

Executive summary

Part 1. What do we now know about sustained attention and multitasking based on the behavioral analysis?

The unknown at the outset of this work was if high cognitive load environments would decrease alertness faster or more slowly. One outcome of the work is that stimulating environments seem to help preserve alertness. However, engaging someone with additional tasks to increase their alertness also resulted in losses to attentional orienting, meaning they missed more information from their environment as the demands of the secondary task increased. Thus, it appears that using secondary tasks to “energize” soldiers may provide benefits to alertness if they are engaged in monotonous tasks, but these tasks may decrease attention to the environment. The conclusion is that in cases of the most extreme declines in alertness, a secondary task is beneficial, but generally secondary tasks pose a threat to overall attentiveness.

Part 2. What do we now know about predicting vigilance losses using advanced statistical and mathematical techniques?

We can make predictions at a gross level across time to chart and predict further alertness losses. Momentary, person-by-person prediction is too influenced by noise and individual variation to allow for a brief sample of data for one person to predict their state of alertness. Alertness waxes and wanes in local windows of time very dramatically. But over time, trends can be plotted using multilevel models of momentary derivative estimates. Soldier readiness can be predicted if enough prior data are collected on readiness states and the proper model is used to analyze those data.

Part 3. Additional pilot work holds promise for deploying a readiness measurement device to give commanders a real-time view of the cognitive readiness of their soldiers both in vehicles and on the battlefield.

Our long-term goal is to develop tools to help commanders gauge the alertness and readiness of their soldiers in any environment, from the driver’s cab to the soldier on the ground. We tested a candidate device that could be deployed with individual soldiers and query their alterness in an unobtrusive way (using either light or vibrotactile stimulation). The data from these queries, when aggregated over time and analyzed using some of the models developed in Part 2, can indicate to a commander the overall alertness state of a soldier and a unit at a moment in time and also show the trend in unit alertness and readiness. Work done here with drivers must be extended to the field to realize this goal. But he initial results are very promising and suggest that additional work will be fruitful.

Goal 1: To better understand potential risks and benefits of multitasking on sustained attention.

This section lists the background, method and behavioral results of the project, and explains the significance of the results.

Quick summary: **Method**: Two driving scenarios were employed in this study to investigate the effects of perceptual load— (1) a congested roadway (high perceptual load) and (2) a monotonous roadway (low perceptual load). Three verbal task conditions were used to investigate the effects of a concurrent task when sustained attention was at its lowest: (1) no verbal, (2) simple verbal, (3) complex verbal. **Results**: Drivers, regardless of perceptual load, exhibited improved performance when sustained attention was at its lowest when a secondary task was introduced. However, performance on a visual attention task was poorer for drivers who engaged in a secondary task continuously. **Conclusion:** Introducing a strategic concurrent task can improve performance when sustained attention is at its lowest. However, this benefit carries the cost of reduced visual attention to objects in the periphery.

Section 1: Brief background

Driving on rural highways with little stimulation or variation in the environment is unlikely to inspire any active task engagement. Under such monotonous conditions, a driver is likely to make mental plans of tasks to carry out upon arrival, think about the day's events, or, possibly, even pick up the cell phone to call a friend. In other words, the driver is not actively engaged in the task of driving. The driver simply goes through the motions of having to get to his destination, while attempting to remain vigilant and react to possible hazards such as stray animals that dash across the roadway. The necessity of maintaining vigilance does not occur only under monotonous conditions. Consider a commuter driving on a crowded interstate where traffic flow is congested and unpredictable. Under such conditions, the driver needs to be aware of the driving environment and respond to any potential hazardous situations such as a sudden breaking event or an inconsiderate motorist attempting to jump the queue. In such situations, the driver is likely to be deeply engaged in the driving task while maintaining a watchful eye on road conditions. Both scenarios described here require drivers to maintain vigilance. However, the willingness and ability to focus and sustain attention to the task of driving differs between the scenarios. When a task becomes simple and well-practiced, such as driving under monotonous conditions, attention is likely to shift toward processes unrelated to the primary task. Conversely, greater environmental demands will likely encourage the maintenance of attention toward the primary task of driving.

Task-unrelated thoughts occur automatically in any situation, especially when the task lacks novelty. Monotony and the lack of stimulation create an environment conducive to the development of these wandering thoughts. The ubiquity of such thoughts in our daily life suggests little harm in daydreaming or mind-wandering. In fact, evidence implicates the role of off-task thoughts in the creative experience (Smallwood & Schooler, 2006). While such thoughts may appear to evoke a sense of wonder and excitement, research also suggests that such thoughts are undesirable as they may impose a negative influence on the performance of certain tasks. The wandering mind is not a new concept (see Antrobus, Singer, & Greenberg, 1966). However, it is only in recent times that researchers have become interested in how such wandering thoughts can influence task performance. The field of sustained attention[1] research has become a recent hotbed for the discussion and development of these ideas.

Research in sustained attention focuses on factors that can influence an individual's ability to maintain vigilance. More important is the goal of identifying methods and countermeasures that can aid in the maintenance and optimization of sustained attention. Theories of sustained attention have continually evolved over the past five decades. The most current and dominant view is the resource depletion model. By virtue of its monotonous and repetitive nature, vigilance tasks require substantial effort to reduce

[1] Most researchers use the terms "vigilance" and "sustained attention" interchangeably. The differences are field specific and academic rather than practical (see Oken, Salinsky, & Elsas, 2006). Clinical researchers use the term "sustained attention", while cognitive and human factors researchers have traditionally used the term "vigilance". If any distinction has to be drawn, vigilance is said to be a more specific form of sustained attention. For the sake of clarity, and continuity, this dissertation will use "sustained attention" and "vigilance" interchangeably.

performance decrements (Finomore, Matthews, Shaw, & Warm, 2009; Frankmann & Adams, 1962; Helton et al., 2005; Warm, Parasuraman, & Matthews, 2008). This claim states that the effort needed to maintain vigilance exhausts resources to the point where performance becomes impaired. This view can be summed up by the "paradox of vigilance," where a seemingly simple monitoring task becomes mentally taxing as it continues to demand large amounts of resource to ensure optimal performance (Gaillard, 2008; Warm et al., 2008). In a perfect world, where resources are unlimited, vigilance levels will remain unchanged. This, however, is unlikely the case in reality, hence the development of computer assisted systems that reduce task-related resource demands (see Parasuraman & Wickens, 2008).

A more recent view purports that poor sustained attention is the result of the redirection of resources toward task-unrelated thoughts. This view runs contrary to the dominant view of sustained attention as it claims that vigilance tasks are undemanding. As the repetitive and monotonous nature of vigilance tasks creates an environment of familiarity and simplicity, behaviors gradually become automatic and mindless (Manly, Robertson, Galloway, & Hawkins, 1999; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). Demand for resources decreases substantially when a task is practiced to the point of becoming automatized (Schneider & Schiffrin, 1977). This implies an increase in the availability of resources towards processing of information unrelated to the maintenance of sustained attention, resulting in a decline in task performance. Such a scenario implies that automation can only serve to worsen vigilance. Paradoxically, there is evidence suggesting that increasing the load of a task can counter the redirection of resources toward task-unrelated thoughts (see Forster & Lavie, 2009). When an individual is sufficiently engaged in a task, particularly when the task has an increased load, attention is less likely to become redirected.

The question of whether resources are depleted or redirected toward task-unrelated thoughts continues to be debated (R. Parasuraman, personal communication, July 20, 2011). While there is strong evidence that resources are depleted, the alternative resource redirection model is beginning to gain attention within the domain of vigilance research. Similar to driving on a rural highway, most vigilance tasks have been conducted under monotonous conditions. While such an experimental setup readily tests vigilance, it fails to fully address the increased engagement that individuals exhibit when placed in a demanding environment. This project aimed to investigate the effects of perceptual and cognitive loads on vigilance specifically within a driving environment.

Section 2: Hypotheses and results summaries

There were two primary goals to this study: (1) to investigate the effects of roadway conditions (perceptual load) on the maintenance of sustained attention and driving performance and (2) the effects of a concurrent verbal task load (cognitive load) on driving performance when sustained attention was at its lowest.

Hypothesis 1: Effects of perceptual load on driving performance. While previous research has found that increased loads result in a rapid decline in sustained attention (see Background), there is contrary evidence suggesting otherwise (see Load on performance). Increasing the load of a task demands greater task engagement to ensure ideal performance. More important is the relationship between task engagement and performance on tasks requiring sustained attention (see Matthews, Warm, Reinerman-Jones, et al., 2010). Due to higher perceptual load in congested traffic, driving performance will initially decline more rapidly. However, due to increased task engagement, decline in performance will asymptote at a higher level than drivers in the monotonous driving condition.

Summary of results from Hypothesis 1. While analysis was conducted separately for each driving condition, results indicated a decline in performance as driving time increased. The effect of time was observed in all dependent measures except steering deflections. Performance generally declined in a linear fashion for both driving conditions. However, a quadratic trend in decline was also observed for the congested conditions. This quadratic trend suggested a curvilinear shape to the decline in performance. In addition, the significant pairwise comparisons between time blocks further supported the linear trend in performance decline particularly for the monotonous driving condition. Conversely, the weak pairwise comparisons for the congested condition support the quadratic trend in performance decline.

Hypothesis 2: Effects of cognitive load on driving performance. The introduction of a concurrent verbal task load improved driver performance when vigilance was lowest (see Atchley & Chan, 2011). While the effects were clearly observed, the drive was only conducted on a monotonous roadway, making it unsuitable to make predictions on high perceptual load conditions. Moreover, by only employing a free association verbal task, Atchley and Chan (2011) could not manipulate the demands of the verbal task. Consequently, this study manipulated the cognitive demands of the concurrent verbal load. Introducing the simple verbal task under congested driving conditions will result in little change to driving behavior. Engaging in a complex verbal task under congested conditions will result in impaired driving performance as demand for resources increases. Having previously reported the benefits of a free association verbal task in monotonous driving conditions, the use of the simple verbal task will likely exhibit a similar level of performance as Atchley and Chan (2011). However, the increased task engagement required by the complex verbal task may result in greater performance benefits under monotonous driving conditions.

Summary of results from Hypothesis 2. In general, the introduction of a verbal task resulted in improved driving performance for both conditions. Engaging in either a simple or complex verbal task resulted in improved performance for drivers in the monotonous condition. However, improved performance for drivers in the congested condition was only observed with a complex verbal task. In addition, the performance improvement resulted in better driving performance than not engaging in a verbal task. This was particularly the case for drivers in the monotonous condition. The benefit of engaging in a verbal task over not having one when driving under congested conditions was only observed under the complex verbal task condition.

In terms of task engagement, drivers in the monotonous condition who did not have a verbal task reported lowest levels of engagement. Conversely, drivers in the congested condition as well as drivers in the monotonous condition, who engaged in a late verbal task, generally reported higher task engagement. While the results suggested better performance, billboard recall was significantly reduced in the presence of increased load.

Section 3: Method

The study employed a simulated drive either on a congested roadway or a monotonous and open environment. Participants for the study were recruited via the University of Kansas undergraduate research pool. Only individuals who had a valid driver's license were selected. Additionally, because of the language requirements of the verbal task, participants were limited to native English speakers only. Ninety-eight individuals from the University of Kansas Undergraduate pool participated in the study. Seven participants were excluded from the analysis for failing to follow instructions, and one participant did not have a valid driver's license. Only 90 participants (termed "drivers" from this point onward) were included in the analysis. There were 49 female and 41 male drivers. Mean age of drivers was 19.2 years (SD = 1.60), and mean driving experience was 4.19 years (SD = 1.95). 32 additional participants were recruited to serve as controls for the verbal task load. There were 11 female and 21 male participants. Mean age of control participants was 19.3 (SD = .82). Control participants did not drive in the simulator.

Materials and apparatus

Driving simulator. The simulator suite was a fixed based unit complete with pedals and force feedback steering. The steering wheel included additional input buttons to capture driver responses to secondary tasks while driving. The driving scene was presented on a single 20" LCD monitor that provided a 45 degree driver field of view. Participants drove on a simulated roadway designed using STISIM Drive (2.08.06). The software provided various measures of driving performance including vehicle position on the roadway, steering deviation, and vehicle velocity. Output from the simulator was collected at 30Hz.

Driving scenarios. There were two driving scenarios in this study. The first scenario was a low load driving condition with little traffic. This scenario was similar to previous studies investigating the effects of roadway monotony on driving performance. The roadway was a four-lane highway, with two inbound and two outbound lanes. Lane width was set at 12 feet and traffic was divided by a 20 feet median. The second scenario was a high load condition, designed to mimic a congested highway in an urban environment undergoing construction, resulting in slow traffic. Physical characteristics of the road such as lane markers were identical in both conditions. Overall density of traffic on the roadway for both scenarios was determined by the following formula:

$$D_F = \frac{\sum_{i=1}^{n} D_i \times L_i \times N_i}{\sum_{i=1}^{n} L_i \times N_i}$$

where

D_F = average density for the roadway (passenger car / mile / lane),
D_i = density for segment *i* (passenger car / mi / lane),
L_i = length of segment *i* (ft),
N = number of lanes in segment *i*, and
n = number of segments in the defined roadway.

Table 4.
Level of Service (LOS) and associated roadway density.

Level of Service	Density (passenger car / mile / lane)
A	≤ 11
B	> 11 - 18
C	> 18 - 26
D	> 26 - 35
E	> 35 -45
F	> 45

Table 4 provides the density (D_F) of the roadway and the associated Level of Service (LOS). The LOS for the monotonous drive had an "A" rating, indicating stable traffic flow and optimal driving speeds. This scenario lasted 45 minutes with no breaks. Drivers drove at approximately 70 to 75 mph. All drivers assigned to this condition were instructed to use the cruise control once the appropriate highway speed was achieved.

The LOS for the congested roadway was rated between "E" and "F", indicating a roadway where traffic flow was suboptimal and unstable. This scenario lasted 45 minutes with no breaks. To adhere to, and mimic congested driving conditions, driving speeds in this condition did not exceed 35 mph. Traffic conditions were designed in accordance to the Highway Capacity Manual (Transportation Research Board, 2010). Physical parameters of the roadway including length of lane dividing lines, lane widths, and lane markers were set in accordance to the Manual on Uniform Traffic Control Devices (Federal Highway Administration, 2009). Both driving scenarios occurred under clear weather conditions with visibility set at 2500 ft.

Visual attention and memory. A recall task of the billboards was administered at the end of the drive to assess the visual attention of drivers. Distracted drivers are more likely to exhibit poorer visual attention (Atchley & Dressel, 2004), and therefore more likely to miss objects in the periphery. Brand logos of common fast food chain restaurants found in North America were used (see Appendix B for logos.). Billboards were located six feet away from the right shoulder. The presentation of billboards was randomized across participants. This served as a measure of load on visual attention. Participants were not informed of the memory task to avoid elevated attention to the billboards. Billboard presentation occurred only in the first and last blocks of the drive and there were four billboards in each block.

Peripheral Detection Task (PDT). The driving simulator software provided visual cues in the periphery of the driving scene to measure a driver's attention to objects in the periphery. The PDT required drivers to make an appropriate button press response on the steering wheel when the symbol changed from a diamond (as shown) to a leftward or rightward pointing arrow. Arrow directions were congruent with the spatial location of the cue, i.e. left diamond would change to a leftward pointing arrow, and right diamond would change to a rightward pointing arrow. Responses to the arrows were similarly mapped on the steering wheel, i.e. left button press for leftward arrows and right button press for rightward arrows. Drivers had five seconds to make a response. All PDT events occurred in the final minute of a driving block. Appendix C provides a schematic and detailed description of driving scenarios, billboards, and secondary task events. Figure 4 provides an example of both driving environments along with the PDT symbols and billboards.

Verbal task load. A modified verb generation task designed by Snyder et al. (2011) was used as the concurrent verbal task (See Appendix D for the wordlist). The task required participants to generate an associated verb response when presented with a particular noun (e.g. "dog"). An appropriate verb response in this case could be associated to the noun in terms of what it does (e.g. "bark"), or an action (e.g. "walk"). The words were normed based on the level of selection and retrieval demands (high vs. low). For the purposes of this study, the words used were divided based on retrieval demands (associations) and collapsed across selection demands. To reduce possible confounds, words normed on the level of selection were evenly divided. The complex verbal task contained nouns with few associated verbs (high retrieval demand) while the simple verbal task (low retrieval demand) contained nouns with multiple associated verbs. The verb generation task began at the start of the final block, approximately 37 minutes into the drive. Words were randomly selected without replacement and presented at an inter-stimulus interval (ISI) of eight seconds. The wordlist was reset upon completion and was randomized again until the driving scenario was terminated. Drivers and control participants had approximately eight seconds to respond to the cue word. Words were pre-recorded by a female native speaker of English and played back via E-prime to ensure consistency. In addition, all auditory stimuli were presented to the right ear via hands-free headset. Responses were recorded for future analysis.

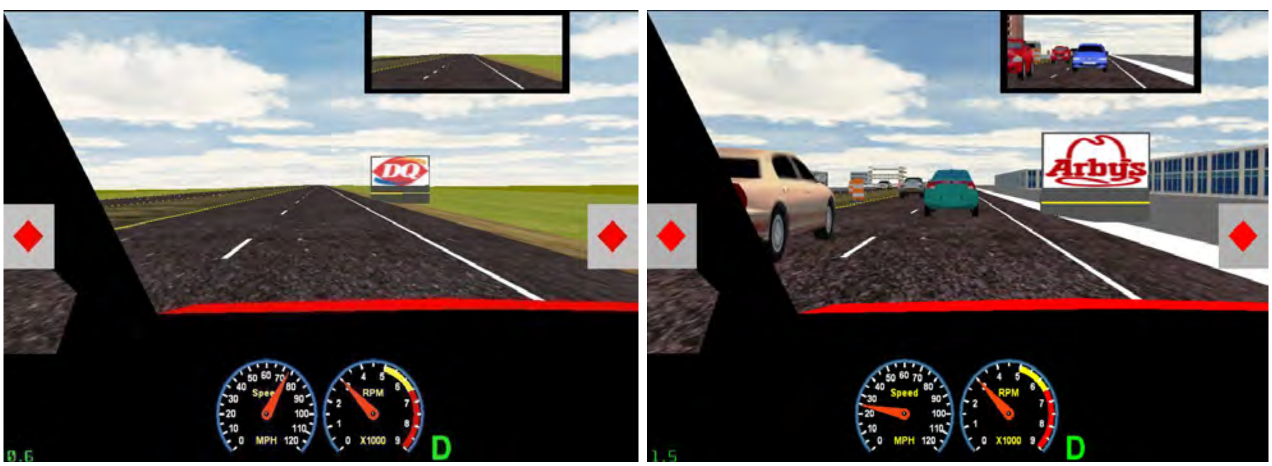


Figure 4. Examples of monotonous (left) and congested (right) roadways. Diamonds are visual cues used in the peripheral detection task.

Workload and task engagement. Task load was assessed post task using the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988). To assess task engagement, the researcher asked drivers and control participants to complete the task engagement section of the Dundee Stress State Questionnaire (DSSQ; Matthews et al., 2002; Matthews, Warm, Reinerman-Jones, et al., 2010) pre and post driving.

Experimental design and procedure

Drivers were randomly assigned into one of six conditions, using a 2 (driving load: monotonous vs. congested) x 3 (verbal task load: no verbal, simple verbal, complex verbal) mixed design. Time was a repeated measure (five time blocks, each nine minutes in duration). Upon informed consent, drivers completed the pre-drive portion of the DSSQ followed by an operation span task[2] (OSPAN; Unsworth, Heitz, Schrock, & Engle, 2005). Once pre-drive measures were complete, drivers had a five-minute practice drive to familiarize themselves with the simulator and the PDT events. Drivers were reminded to observe safe driving behaviors such as: speed adherence, maintenance of safe headway distance, and use of indicators where appropriate. Drivers were reminded to remain in the right lane, unless passing another vehicle. In addition, drivers assigned to the monotonous driving condition were instructed to engage the cruise control system once they achieved highway speeds (70 to 75 mph). All drivers regardless of verbal task load condition practiced the verbal task. Practice words were not used in the actual study. Drivers were reminded to turn their cell phones OFF and were also asked to remove their watches to reduce possible end of task performance bursts. A researcher sat in an adjoining room to observe participants via close-circuit television to ensure compliance and appropriate driving behavior and trigger events during the drive. Upon completion of the driving scenario, drivers were given three minutes to freely recall the billboards they saw on the roadway. Drivers then completed the second half of the DSSQ along with the MMI[3] (Ophir, Nass, & Wagner, 2009) and an abbreviated version of the NASA-TLX. Total duration of the experiment was approximately 100 minutes.

Participants in the control condition were randomly assigned into one of two verbal task loads (simple vs. complex). Participants were first asked to complete a pre-task section of the DSSQ to assess task engagement. Participants then proceeded with the verbal task (nine minutes) while passively sitting in the driving simulator. Upon completing the verbal task, control participants proceeded to complete the second half of the DSSQ to assess post-task engagement scores. The control task lasted 20 minutes.
The assessment of performance in most vigilance tasks often rely on signal detection theories and other psycho-physical measures. However, the dynamic nature of the driving environment will require other proxy measures of sustained attention. The following section will outline the measures used in this project investigate driving performance.

Performance measures

Weaving & lane infractions. Weaving was measured by the standard deviation of lane position (SDLP). Larger SDLP values indicated a decline in performance with prolonged time on task. Increases in SDLP are attributed to monotonous driving conditions which results in decreased attention (Desmond & Matthews, 1997). Greater SDLP values

[2] The OSPAN measure was collected for an unrelated investigation and will not be further discussed in the dissertation.
[3] The MMI measure was collected for an unrelated investigation and will not be discussed in the dissertation.

are also associated with an increase in collisions with vehicles in adjacent lanes (Ranney, Harbluk, & Noy, 2005).

The root mean square error (RMSE) of lane position was another measure for investigating lateral stability while driving. While SDLP measures deviation from the driver's individual mean, RMSE measures a driver's deviation from a predefined midline within the lane. Both measures (SDLP and RMSE) remain in use within the literature (e.g. Drews, Pasupathi, & Strayer, 2008). Lane infractions were defined by the following instances: crossing into an adjacent lane when there were no vehicles to pass and driving in the median or road shoulder. Reductions in attention of the drivers will likely result in more infractions (Liu & Wu, 2009).

Steering wheel angle & steering deflections. A driver has to make minor steering adjustments while driving to ensure steady lane keeping. An increase in the variability of steering wheel angle indicated a reduction in attention (Brookhuis & De Waard, 1993; Oron-Gilad & Ronen, 2007; Thiffault & Bergeron, 2003). Rather than using the standard deviation of steering wheel angle which varied widely between drivers, the root mean square error (RMSE) of the steering angle was used instead. The RMSE value measures the deviation of steering wheel angle from zero degrees. This was a more accurate measure as the steering wheel was centered at zero degrees for all drivers. Thiffault and Bergeron (2003) also suggested that steering deflections greater than 10 degrees indicate a reduction in attention as drivers are making sharp adjustments rather than minor ones to keep the vehicle on course.

Workload, task engagement, and attention. Scores on the NASA-TLX were used to assess perceived driver workload, and scores from the DSSQ were used to assess task engagement. To measure attention to the roadway, participants were required to recall as many billboards they remember seeing on the roadway. This value was changed to a proportion during analysis. Reaction time and accuracy to the PDT were used as additional measures of attention.

Section 4: Results

Hypothesis 1: Perceptual load and driving performance.

This portion of the analysis focused on the effect of time on performance under different driving loads. Only data from time block one to four were used. Data from time block five was excluded due to the presence of concurrent verbal tasks. Only data from the right hand lane were used as drivers were instructed to remain primarily in the right lane unless passing another vehicle. Due to systematic differences in vehicle speed, analysis only focused on performance within each driving load condition. Unless stated otherwise, repeated measures ANOVA were used to investigate changes in performance for each driving load. Violations of sphericity were corrected using Greenhouse-Geisser or Huynh-Feldt adjustments when epsilon (ε) was < 0.70 or ≥ 0.70 (Howell, 2002) respectively. Bonferroni-Holm adjustments were applied to multiple comparisons, and null-hypothesis significance tests were conducted at $\alpha = .05$.

Standard Deviation of Lane Position (SDLP)

Monotonous drive. Results indicated a significant ($\varepsilon = .68$) reduction in lane keeping performance over time, $F(2.06, 90.79) = 37.33$, $p < .001$, $\eta^2_p = .459$. Drivers exhibited a linear decline in performance, $F(1, 44) = 60.03$, $p < .001$, $\eta^2_p = .577$. Both quadratic, $F(1, 44) = 3.09$, $p = .085$, $\eta^2_p = .066$, and cubic trends, $F(1, 44) = 3.39$, $p = .072$, $\eta^2_p = .071$, were not significant. Pairwise comparisons revealed significant reduction in performance between time block one and time block two, $t(44) = - 7.30$, $p < .001$, $d = - 0.535$; time block two and time block three, $t(44) = - 3.23$,.$p = .005$, $d = - 0.208$, and time block three and time block four , $t(44) = - 2.79$, $p = .008$, $d = - 0.243$ (see Table 5)

Congested drive. Results indicated a significant ($\varepsilon = .84$) reduction in lane keeping performance over time, $F(2.51, 104.20) = 25.77$, $p < .001$, $\eta^2_p = .369$. Drivers exhibited both linear, $F(1, 44) = 46.59$, $p < .001$, $\eta^2_p = .514$, and quadratic trends, $F(1, 44) = 4.24$, $p = .045$, $\eta^2_p = .088$, in performance declines. The cubic trend was not significant, $F(1, 44) = .1$, $p = .753$, $\eta^2_p = .002$. Pairwise comparisons revealed significant reduction in performance between time block one and time block two, $t(44) = - 4.23$, $p < .001$, $d = - 0.428$; time block two and time block three, $t(44) = - 3.75$, $p < .001$, $d = - 0.255$, but not between time block three to time block four, $t(44) = - 1.36$, $p = .182$, $d = - 0.090$ (see Table 5).

Root Mean Square Error (RMSE) – Lane Position

Monotonous drive. Results indicated a significant ($\varepsilon = .79$) reduction in lane keeping performance over time, $F(2.38, 104.55) = 31.35$, $p < .001$, $\eta^2_p = .416$. Drivers exhibited both linear, $F(1, 44) = 46.12$, $p < .001$, $\eta^2_p = .512$, and quadratic trends, $F(1, 44) = 16.39$, $p < .001$, $\eta^2_p = .271$, in performance decline. The cubic trend was not significant, $F(1, 44) = 1.09$, $p = .302$, $\eta^2_p = .024$. Pairwise comparisons revealed a reduction in performance between time block one and time block two, $t(44) = -7.79$, $p < .001$, $d = - 0.528$; time block two and time block three, $t(44) = - 2.53$, $p = .030$, $d = - 0.185$, but not between time block three and time block four, $t(44) = - 1.02$, $p = .312$, $d = - 0.076$ (see Table 5).

Congested drive. Results indicated a significant ($\varepsilon = .76$) reduction in lane keeping performance over time, $F(2.29, 100.81) = 17.49$, $p < .001$, $\eta^2_p = .284$. Drivers exhibited both linear, $F(1, 44) = 27.19$, $p < .001$, $\eta^2_p = .382$, and quadratic trends, $F(1, 44) = 13.69$, $p < .001$, $\eta^2_p = .237$, in performance declines. The cubic trend was not significant, $F(1, 44) = 2.35$, $p = .132$, $\eta^2_p = .051$. Pairwise comparisons revealed significant reduction in performance between time block one and time block two, $t(44) = -7.50$, $p < .001$, $d = -0.576$. There was no significant change in performance decline between time block two and time block three, $t(44) = -0.72$, $p = .950$, $d = -0.073$, and time block three and time block four, $t(44) = -0.65$, $p = .950$, $d = -0.062$ (see Table 5).

Table 5[a]
Standard deviation of lane position (SDLP) and root mean square error (RMSE) of lane position. All values are in feet (ft).

	Time[b] Block 1	Block 2	Block 3	Block 4
Roadway	SDLP (ft)			
Monotonous	1.02 (.20)	1.14 (.23)	1.19 (.25)	1.26 (.29)
Congested	0.65 (.17)	0.73 (.22)	0.79 (.25)	0.81 (.25)
Roadway	RMSE of Lane Position (ft)			
Monotonous	1.09 (.28)	1.25 (.29)	1.30 (.32)	1.33 (.32)
Congested	1.07 (.39)	1.30 (.41)	1.33 (.49)	1.36 (.41)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Root mean square error (RMSE) - Steering Deviation

Monotonous drive. Results indicated a significant ($\varepsilon = .72$) increase in steering deviation over time, $F(2.15, 94.72) = 25.85$, $p < .001$, $\eta^2_p = .370$. Drivers exhibited a linear decline in performance, $F(1, 44) = 45.86$, $p < .001$, $\eta^2_p = .510$. There was no significant quadratic, $F(1, 44) = 1.35$, $p = .252$, $\eta^2_p = .030$, or cubic trend, $F(1, 44) = 0.90$, $p = .348$, $\eta^2_p = .020$, in performance decline. Pairwise comparisons revealed significant reduction in performance between time block one and time block two, $t(44) = -5.71$, $p < .001$, $d = -0.462$; time block two and time block three, $t(44) = -3.83$, $p < .001$, $d = -0.358$, but not between time block three and time block four, $t(44) = -1.51$, $p = .138$, $d = -0.138$ (see Table 6).

Congested drive. Results indicated a significant ($\varepsilon = .90$) increase in steering deviation over time, $F(2.70, 118.89) = 7.13$, $p < .001$, $\eta^2_p = .139$. Drivers exhibited a linear decline in performance, $F(1, 44) = 17.05$, $p < .001$, $\eta^2_p = .279$. There was no significant quadratic, $F(1, 44) = .006$, $p = .938$, $\eta^2_p = .000$, or cubic trend, $F(1, 44) = 1.07$, $p = .307$, $\eta^2_p = .024$, in performance decline. Pairwise comparisons revealed a marginal reduction in steering deviation between time block two and time block three, $t(44) = -2.48$, $p = .051$, $d = -0.313$, and no significant reduction in performance between time block one and time block two, $t(44) = -1.06$, $p = .594$, $d = -0.179$; and time block three and time block four, $t(44) = -0.95$, $p = .594$, $d = -0.126$ (see Table 6).

Table 6[a]
Root mean square error (RMSE) of steering wheel deviation (in degrees) and steering wheel deflections greater than 10˚.

	Time[b] Block 1	Block 2	Block 3	Block 4
Roadway	RMSE of steering wheel deviation			
Monotonous	0.73 (.29)	0.88 (.38)	1.05 (.59)	1.13 (.58)
Congested	1.15 (.47)	1.25 (.63)	1.56 (.99)	1.62 (.81)
Roadway	Steering deflections > 10˚			
Monotonous	0.18 (.44)	0.36 (.83)	1.98 (6.98)	1.38 (3.48)
Congested	1.27 (1.63)	1.09 (2.57)	3.24 (10.84)	2.47 (3.75)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Steering Deflections (> 10˚)

Monotonous drive. Results did not indicate a significant ($\varepsilon = .48$) change in steering deflections over time, $F(1.46, 27.58) = 2.45$, $p = .109$, $\eta^2_p = .053$. While not significant, change in performance reflected a significant linear trend, $F(1, 44) = 7.37$, $p = .009$, $\eta^2_p = .143$. There was no significant quadratic, $F(1, 44) = 0.54$, $p = .468$, $\eta^2_p = .012$, or cubic trend, $F(1, 44) = 1.58$, $p = .215$, $\eta^2_p = .035$, in performance decline. Pairwise comparisons did not reveal significant changes in performance between time block one and time block two, $t(44) = -1.67$, $p = .309$, $d = -0.279$; time block two and time block three, $t(44) = -1.66$, $p = .309$, $d = -0.415$, and time block three and time block four, $t(44) = 0.57$, $p = .569$, $d = 0.115$ (see Table 6).

Congested drive. Results did not indicate a significant ($\varepsilon = .41$) change in steering deflections over time, $F(1.22, 53.85) = 1.69$, $p = .200$, $\eta^2_p = .037$. While not significant, change in performance reflected a significant linear trend, $F(1, 44) = 4.74$, $p = .035$, $\eta^2_p = .097$. There was no significant quadratic, $F(1, 44) = .14$, $p = .710$, $\eta^2_p = .003$, or cubic trend, $F(1, 44) = 1.62$, $p = .210$, $\eta^2_p = .035$, in performance decline. Pairwise comparisons did not reveal changes in performance between time block one and time block two, $t(44) = 0.45$, $p = 1$, $d = 0.085$; time block two and time block three, $t(44) = -1.49$, $p = .433$, $d = -0.321$, and time block three and time block four, $t(44) = 0.55$, $p = 1$, $d = 0.107$ (see Table 6).

Lane infraction count

Monotonous drive. Results indicated a significant (ε = .57) increase in lane infractions over time, $F(1.70, 74.98) = 13.74$, $p < .001$, $\eta^2_p = .238$. Drivers exhibited a linear decline in performance, $F(1, 44) = 20.86$, $p < .001$, $\eta^2_p = .322$. Both quadratic, $F(1, 44) = 0.05$, $p = .832$, $\eta^2_p = .001$, and cubic trends, $F(1, 44) = 0.75$, $p = .392$, $\eta^2_p = .017$, were not significant. Pairwise comparisons revealed significant reductions in performance between time block one and time block two, $t(44) = - 3.25$, $p = .003$, $d = - 0.468$, and time block two and time block three, $t(44) = - 3.59$, $p = .004$, $d = - 0.423$. The number of infractions between time block three and time block four were not significant, $t(44) = 1.69$, $p = .096$, $d = - 0.213$ (see Table 7).

Congested drive. Results indicated a significant (ε = .89) increase in lane infractions over time, $F(2.67, 117.42) = 6.36$, $p < .001$, $\eta^2_p = .126$. Drivers exhibited both linear, $F(1, 44) = 12.72$, $p < .001$, $\eta^2_p = .224$, and quadratic trends, $F(1, 44) = 7.33$, $p = .009$, $\eta^2_p = .143$, in performance decline. The cubic trend was not significant, $F(1, 44) = .88$, $p = .353$, $\eta^2_p = .020$. Pairwise comparisons did not reveal a significant change in lane infractions between time block one and time block two, $t(44) = - 2.28$, $p = .083$, $d = - 0.424$; time block two and time block three, $t(44) = - 1.76$, $p = .172$, $d = - 0.217$, and time block three and time block four, $t(44) = 1.75$, $p = .172$, $d = 0.202$ (see Table 7).

Table 7[a]

Number of lane infractions and duration of lane infractions in milliseconds (msec)

	Time[b] Block 1	Block 2	Block 3	Block 4
Roadway	Lane infractions			
Monotonous	0.84 (1.39)	1.82 (2.78)	3.38 (4.56)	4.51 (6.08)
Congested	0.18 (.53)	0.67 (1.77)	1.07 (1.92)	0.73 (1.37)
Roadway	Duration of lane infractions (msec)			
Monotonous	1106 (2424)	2404 (4183)	4796 (6932)	6706 (8944)
Congested	343 (1000)	1226 (3743)	2043 (3513)	1286 (2571)

[a] Numbers in parentheses are standard deviations

\ [b] Each time block is 9 mins in duration

Lane infraction duration

Monotonous drive. Results indicated (ε = .58) longer infractions over time, $F(1.75, 76.98) = 15.53$, $p < .001$, $\eta^2_p = .261$. Drivers exhibited a linear decline in performance, $F(1, 44) = 23.67$, $p < .001$, $\eta^2_p = .350$. Both quadratic, $F(1, 44) = 0.32$, $p = .576$, $\eta^2_p = .007$, and cubic trends, $F(1, 44) = 0.97$, $p = .329$, $\eta^2_p = .022$, were not significant. Pairwise comparisons revealed a significant increase in infraction duration between time block one and time block two, $t(44) = - 2.79$, $p = .015$, $d = - 0.393$; time block two and time block three, $t(44) = - 3.77$, $p < .001$, $d = - 0.430$, and time block three and time block four, $t(44) = - 2.05$ $p = .046$, $d = - 0.241$ (see Table 7).

Congested drive. Results indicated (ε = .91) increased infraction duration over time, $F(2.73, 120.29) = 5.98$, $p < .001$, $\eta^2_p = .120$. Drivers exhibited linear, $F(1, 44) = 14.25$, $p < .001$, $\eta^2_p = .245$, and quadratic trends, $F(1, 44) = 6.98$, $p = .011$, $\eta^2_p = .137$, in performance decline. The cubic trend was not significant, $F(1, 44) = 1.14$, $p = .292$, $\eta^2_p = .025$. Pairwise comparisons did not reveal a significant change in duration of infractions between time block one and time block two, $t(44) = - 1.88$, $p = .150$, $d = - 0.372$; time block two and time block three, $t(44) = - 1.85$, $p = .150$, $d = - 0.225$, and time block three and time block four, $t(44) = 2.01$, $p = .150$, $d = 0.249$ (see Table 7).

Combined analysis between driving conditions

The preceding analyses on driving measures were separated by driving load as differences in vehicle speed made comparisons difficult to interpret. However, the following analyses investigated non-driving related measures and were suitable for analysis as there was no confound of speed.

Attention – Peripheral detection task (PDT)

Chi-square tests conducted at each time block comparing the relationship between responses and driving load did not reveal a significant pattern of results. Time block 1: $\chi^2(1, N = 45) = 1.11$, $p = .485$, Cramer's $V = .111$; Time block 2: $\chi^2(1, N = 45) = 3.55$, $p = .104$, Cramer's $V = .199$; Time block 3: $\chi^2(1, N = 45) = 3.46$, $p = .118$, Cramer's $V = .196$; Time block 4: $\chi^2(1, N = 45) = 2.25$, $p = .230$, Cramer's $V = .158$. Responses to the PDT did not differ between driving loads.

Attention – Response time to peripheral detection task (PDT). There was a significant main effect of time, $F(3, 168) = 5.89$, $p < .001$, $\eta^2_p = .095$. There was neither a main effect of driving load, $F(1, 56) = .259$, $p = .613$, $\eta^2_p = .005$, nor a significant time by driving load interaction, $F(3, 168) = 1.52$, $p = .212$, $\eta^2_p = .026$. Reaction time to the PDT slowed down over time (see Table 8). Three, 2 (Driving load) x 2 (Time block) mixed factorial ANOVAs were run to compare changes between successive time blocks.

Time blocks one and two. There was a significant main effect of time, $F(1, 68) = 21.53$, $p < .001$, $\eta^2_p = .240$, and a significant time by drive load interaction, $F(1, 68) = 5.14$, $p = .027$, $\eta^2_p = .070$. There was no effect of drive load, $F(1, 68) = 0.43$, $p = .513$, $\eta^2_p = .006$. Investigation of the interaction was conducted by testing the effect of time on each level of driving load, and the effect of driving load at each time block.

Effect of time on driving load. Drivers in the monotonous driving condition exhibited slower reaction time, $t(36) = - 2.08$, $p = .045$, $d = - 0.415$, between time block one (M = 1252 msec, SD = 755) and time block two (1581 msec, SD = 829). Drivers in the congested condition similarly exhibited a decline in reaction time, $t(32) = -4.09$, $p < .001$, $d = -1.129$, between time block one (M = 1039 msec, SD = 449) and time block two (M =1997 msec, SD = 1248).

Effect of driving load at each time block. There was no significant difference between monotonous driving (M = 1252 msec, SD = 755), and congested driving (M = 1039 msec, SD = 449) at time block one, $F(1, 68) = 2.00$, $p = .162$, $\eta^2_p = .029$. Similarly, there was no difference between monotonous driving (1581 msec, SD = 829), and congested driving (M =1997 msec, SD = 1248) at time block two, $F(1, 68) = 2.73$, $p = .102$, $\eta^2_p = .039$. The ordinal nature of the interaction did not provide additional information beyond the main effect that reaction time slowed down over time.

Time blocks two and three. There was no main effect of time, $F(1, 64) = 3.10$, $p = .083$, $\eta^2_p = .046$, and driving load, $F(1, 64) = 2.887$, $p = .094$, $\eta^2_p = .043$. The time by driving load interaction was not significant, $F(1, 64) = .001$, $p = .981$, $\eta^2_p = .000$.

Time blocks three and four. There was no main effect of time, $F(1, 67) = 0.84$, $p = .364$, $\eta^2_p = .012$, and driving load, $F(1, 67) = .665$, $p = .418$, $\eta^2_p = .010$. The time by driving load interaction was not significant, $F(1, 67) = .703$, $p = .405$, $\eta^2_p = .010$.

Table 8[a]

Reaction time in milliseconds (msec) to peripheral detection task (PDT).

	Time Block 1	Block 2	Block 3	Block 4
Roadway	Reaction time (msec)			
Monotonous	1262 (799)	1585 (861)	1367 (736)	1540 (890)
Congested	966 (331)	1872 (1310)	1613 (1228)	1612 (844)

[a] Numbers in parentheses are standard deviations

[b] Each time block is 9 mins in duration

Hypothesis 2: Effects of cognitive load on driving performance

Due to systematic differences in vehicle speed, analysis on performance was only conducted within driving loads. Unless stated otherwise, a 3 (Verbal task load) x 2 (Time) Mixed ANOVA was used to analyze the effects of verbal task load at the last two time blocks for each driving condition. A priori predictions based on previous research by Atchley and Chan (2011) warranted analysis to investigate the effects of verbal load between time blocks four and five. Additional analysis was conducted at time block five to investigate differences between verbal loads. Bonferroni-Holm adjustments were applied to multiple comparisons and all null-hypothesis significance tests were conducted at $\alpha = .05$.

Standard Deviation of Lane Position (SDLP)

Monotonous drive. There was a significant main effect of time, $F(1, 42) = 59.90$, $p < .001$, $\eta^2_p = .588$, and verbal load, $F(2, 42) = 8.16$, $p < .001$, $\eta^2_p = .280$. The time by verbal task load interaction was significant, $F(2, 42) = 16.06$, $p < .001$, $\eta^2_p = .433$ (see Figure 5). The interaction was investigated by testing the effect of time on each level of verbal task load and the effect of verbal task load at each time block.

Effect of time on each level of verbal task load. There was no significant change in lane keeping performance between time blocks four and five when drivers did not have a verbal task load, $t(14) = 0.46$, $p = .654$, $d = .041$. However, the introduction of a simple verbal load, $t(14) = 4.08$, $p < .001$, $d = 0.821$, or a complex verbal load, $t(14) = 7.94$, $p < .001$ $d = 1.561$, at time block five resulted in improved lane keeping (see Table 9).

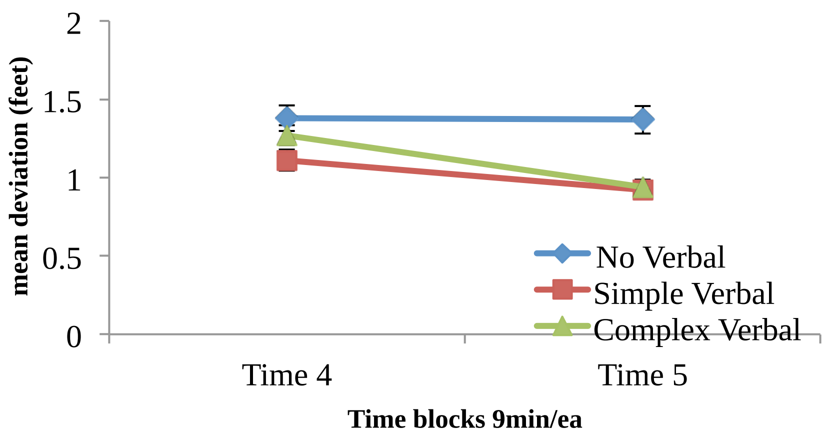


Figure 5[a]. Time by verbal task load interaction for standard deviation of lane position (SDLP) under monotonous conditions. Bars represent standard error of the mean.

Effect of verbal task load at each time block. Comparisons at time block four revealed no difference in lane keeping performance between no verbal and complex verbal load, $t(28) = 1.02$, $p = .319$, $d = 0.374$, as well as, simple verbal and complex verbal loads, $t(28) = -1.82$, $p = .158$ $d = -0.666$. However, lane keeping performance for the no verbal load group was significantly different from the simple verbal load group, $t(28) = 2.56$, $p = .048$, $d = 0.940$ (see Table 9).

Comparisons at time block five revealed that drivers who engaged in a simple verbal load, $t(28) = 4.37$, $p < .001$, $d = 1.652$, or a complex verbal load, $t(28) = 4.22$, $p < .001$, $d = 1.609$, exhibited better lane keeping performance than drivers who had no verbal load. There was no difference in performance between simple and complex verbal loads, $t(28) = -.33$, $p = .745$, $d = -0.119$ (see Table 9).

Table 9[a]
Standard deviation of lane position (SDLP) between verbal task loads at time blocks four and five under monotonous driving conditions. All values are in feet (ft).

	Time[b]	
	Block 4	Block 5
Verbal load	SDLP (ft)	
No verbal load	1.38 (.32)	1.37 (.34)
Simple verbal load	1.11 (.26)	0.92 (.19)
Complex verbal load	1.27 (.24)	0.94 (.18)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Congested drive. There was a significant main effect of time, $F(1, 42) = 11.85$, $p < .001$, $\eta^2_p = .220$. There was no effect of verbal load, $F(2, 42) = 0.59$, $p = .560$, $\eta^2_p = .027$. The time by verbal task load interaction was significant, $F(2, 42) = 10.82$, $p < .001$, $\eta^2_p = .340$ (see Figure 6). The interaction was investigated by testing the effect of time on each level of verbal task load, and the effect of verbal task load at each time block.

Effect of time on each level of verbal load. There was no significant change in lane keeping behavior performance between time blocks four and five when drivers did not have a verbal task load, $t(14) = -2.02$, $p = .063$, $d = -0.268$. The introduction of a verbal task load at time block five resulted in improved driving performance for both simple verbal, $t(14) = 2.22$, $p = .044$, $d = 0.389$, and complex verbal loads, $t(14) = 4.42$, $p < .001$, $d = 0.837$ (see Table 10).

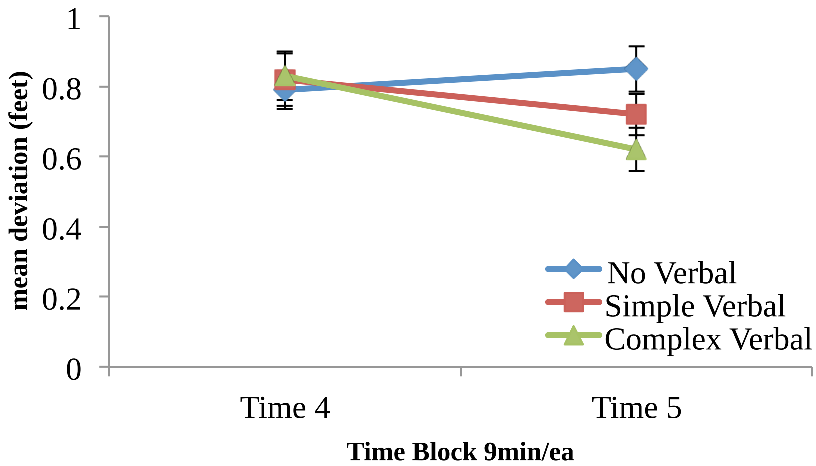


Figure 6[a]. Time by verbal task load interaction for standard deviation of lane position (SDLP) under congested conditions. Bars represent standard error of the mean.

Effect of verbal task at each time block. Comparisons at time block four did not reveal any significant difference in lane keeping performance between no verbal and simple verbal loads, $t(28) = -0.31$, $p = .758$, $d = -0.115$; no verbal and complex verbal loads, $t(28) = -0.49$, $p = .622$, $d = -0.183$; simple verbal and complex verbal loads, $t(28) = -0.15$, $p = .882$, $d = -0.054$ (see Table 10).

Comparisons at time block five revealed that drivers who engaged in a complex verbal load exhibited better lane keeping performance than drivers who did not have a concurrent verbal load, $t(28) = 2.58$, $p = .045$, $d = 0.944$. There was no difference in lane keeping performance between no verbal and simple verbal loads, $t(28) = 1.57$, $p = .276$, $d = 0.558$, as well as, simple verbal and complex verbal loads, $t(28) = 1.15$, $p = .276$, $d = 0.418$ (see Table 10).

Table 10[a]

Standard deviation of lane position (SDLP) between verbal task loads at time block four and five under congested driving conditions. All values are in feet (ft).

	Time[b]	
	Block 4	Block 5
Verbal load	SDLP (ft)	
No verbal load	0.79 (.21)	0.85 (.25)
Simple verbal load	0.82 (.29)	0.72 (.23)
Complex verbal load	0.83 (.27)	0.62 (.24)

[a] Numbers in parentheses are standard deviations

[b] Each time block is 9 mins in duration

Root Mean Square Error (RMSE) – Lane Position

Monotonous drive. There was a significant main effect of time, $F(1, 42) = 28.86$, $p < .001$, $\eta^2_p = .407$, and a significant time by verbal task load interaction, $F(2, 42) = 6.50$, $p = .003$, $\eta^2_p = .236$ (see Figure 7). There was no main effect of verbal task load, $F(2, 42) = 2.69$, $p = .079$, $\eta^2_p = .114$. The interaction was investigated by testing the effect of time on each level of verbal task load, and the effect of verbal task load at each time block.

Effect of time on each level of verbal task load. There was no significant change in lane keeping performance between time blocks four and five when drivers did not engage in a verbal load, $t(14) = 1.18$, $p = .258$, $d = .119$. The introduction of a concurrent verbal load resulted in improved lane keeping performance for drivers in the simple verbal load, $t(14) = 2.63$, $p = .019$, $d = 0.408$, and complex verbal load groups, $t(14) = 4.75$, $p < .001$ $d = 0.824$ (see Table 11).

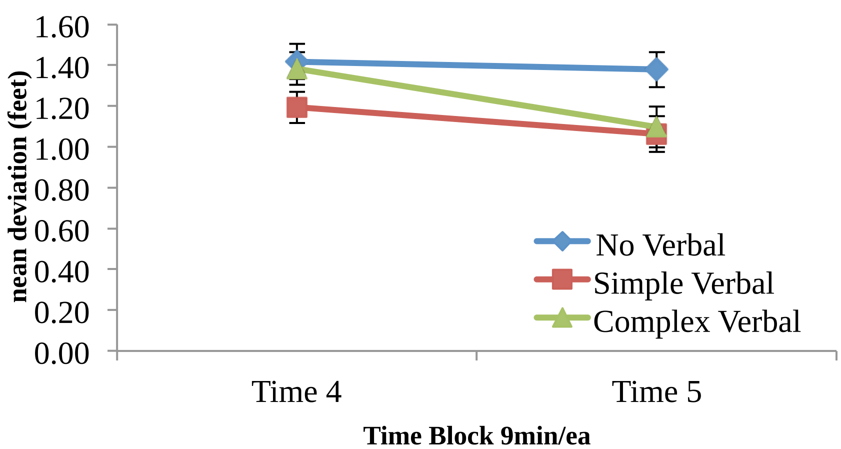


Figure 7[a]. Time by verbal task load interaction for root mean square error (RMSE) – lane position – under monotonous conditions. Bars represent standard error of the mean.

Effect of verbal task load at each time block. Comparisons at time block four did not reveal any significant differences in lane keeping performance between no verbal and simple verbal loads, $t(28) = 1.96$, $p = .060$, $d = 0.718$; no verbal and complex verbal loads, $t(28) = 0.29$, $p = .774$, $d = 0.106$; as well as simple verbal and complex verbal loads, $t(28) = -1.73$, $p = .095$, $d = 0.631$ (see Table 11).

Comparisons at time block five revealed that engaging in a simple verbal load resulted in better lane keeping than not having a verbal load, $t(28) = 2.56$, $p = .048$, $d = 0.936$. There was no difference in lane keeping performance between complex verbal load and no verbal load groups, $t(28) = 2.14$, $p = .083$, $d = 0.783$. There was also no difference in lane keeping performance between simple verbal and complex verbal load groups, $t(28) = -0.26$, $p = .800$, $d = -0.093$ (see Table 11).

Table 11[a]
Root mean square error (RMSE) of lane position between verbal task loads at time blocks four and five under monotonous conditions. All values are in feet (ft)

	Time[b]	
	Block 4	Block 5
Verbal load	RMSE – Lane Position	
No verbal load	1.42 (.33)	1.37 (.33)
Simple verbal load	1.19 (.29)	1.06 (.34)
Complex verbal load	1.38 (.31)	1.09 (.38)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Congested drive. There was a significant main effect of time, $F(1, 42) = 7.05$, $p = .011$, $\eta^2_p = .144$, and a significant time by verbal task load interaction, $F(2, 42) = 4.49$, $p = .017$, $\eta^2_p = .176$. There was no main effect of verbal task load, $F(2, 42) = 0.35$, $p = .709$, $\eta^2_p = .016$. The interaction was investigated by testing the effect of time on each level of verbal task load, and the effect of verbal task load at each time block (see Figure 8).

Effect of time on each level of verbal task load. There was no significant change in lane keeping behavior performance between time blocks four and five when drivers did not engage in a verbal load, $t(14) = 0.63$, $p = .542$, $d = 0.075$, or when drivers engaged in a simple verbal load, $t(14) = 0.15$, $p = .886$, $d = 0.021$. However, the introduction of a complex verbal load resulted in significantly better lane keeping performance, $t(14) = 3.68$, $p = .002$, $d = 0.621$ (see Table 12).

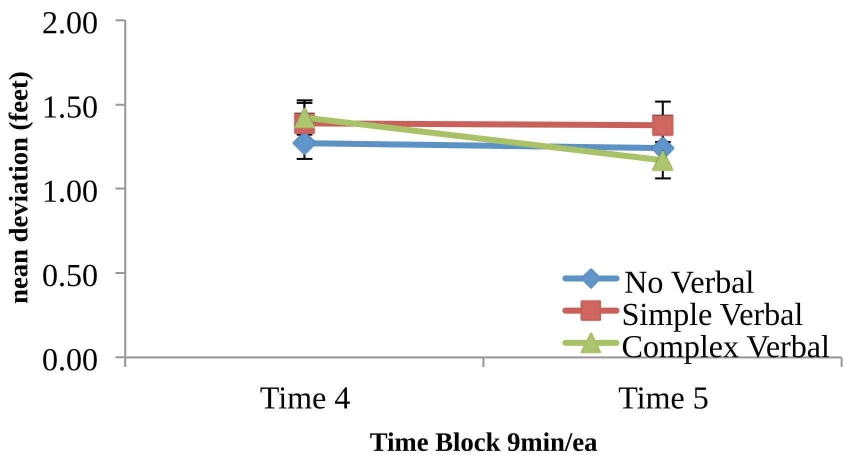


Figure 8[a]. Time by verbal task load interaction for root mean square error (RMSE) – lane position – under congested conditions. Bars represent standard error of the mean.

Effect of verbal task load at each time block. Comparisons at time block four did not reveal any significant differences between no verbal and simple verbal loads, $t(28) = -0.76$, $p = .453$, $d = -0.280$; no verbal and complex verbal loads,: $t(28) = -1.09$, $p = .281$, $d = -0.402$, and simple verbal and complex verbal loads, $t(28) = -0.23$, $p = .824$, $d = -0.082$ (see Table 12).

Table 12[a]

Root mean square error (RMSE) of lane position between verbal task loads at time block four and five under congested conditions. All values are in feet (ft).

	Time[b] Block 4	Block 5
Verbal load	RMSE – Lane Position	
No verbal load	1.27 (.36)	1.24 (.43)
Simple verbal load	1.38 (.47)	1.37 (.55)
Complex verbal load	1.42 (.39)	1.17 (.42)

[a] Numbers in parentheses are standard deviations

[b] Each time block is 9 mins in duration

Comparisons at time block five did not reveal any significant differences between no verbal and simple verbal loads, $t(28) = -0.76$, $p = .456$, $d = -0.278$; no verbal and complex verbal loads,: $t(28) = 0.47$, $p = .645$, $d = 0.170$, and simple verbal and complex verbal loads, $t(28) = 1.16$, $p = .255$, $d = 0.428$ (see Table 12).

Steering Deviation – RMS Steering

Monotonous drive. There was a significant main effect of time, $F(1, 42) = 18.64$, $p < .001$, $\eta^2_p = .307$, and verbal task load, $F(2, 42) = 4.19$, $p = .022$, $\eta^2_p = .166$. There was a significant time by verbal task load interaction, $F(2, 42) = 4.89$, $p = .012$, $\eta^2_p = .189$ (see Figure 9). The interaction was investigated by testing the effect of time on each level of verbal task load, and the effect of verbal task load at each time block.

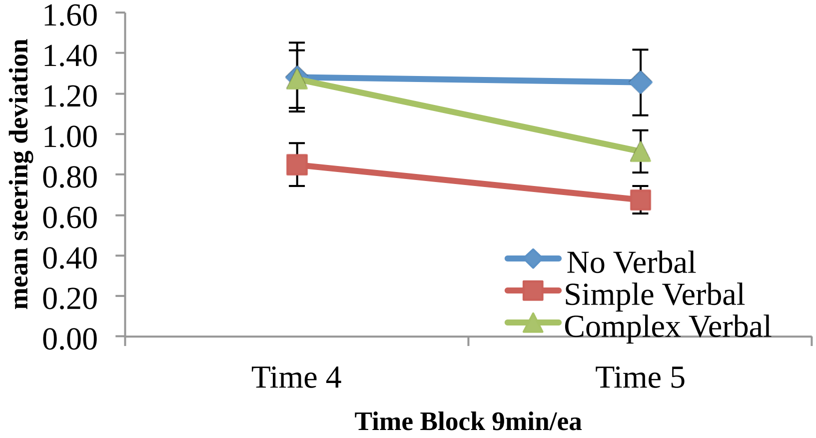


Figure 9[a]. Time by verbal task load interaction for root mean square error (RMSE) – steering – under monotonous conditions. Bars represent standard error of the mean.

Effect of time on each level of verbal task load. There was no significant change in steering deviation between time blocks four and five when drivers did not engage in a verbal task, $t(14) = 1.18$, $p = .258$, $d = 0.043$. However, there was a significant reduction in steering deviation when drivers engaged in a simple verbal task, $t(14) = 2.63$, $p = .019$, $d = 0.513$, or when drivers engaged in a complex verbal task, $t(14) = 4.75$, $p < .001$, $d = 0.744$ (see Table 13).

Effect of verbal task load at each time block. Comparisons at time block four did not reveal significant differences in steering deviation between no verbal and simple verbal loads, $t(28) = 2.16$, $p = .083$, $d = 0.808$; no verbal and complex verbal loads,: $t(28) = 0.49$, $p = .961$, $d = 0.018$; and simple verbal and complex verbal loads, $t(28) = -2.37$, $p = .075$, $d = -0.875$ (see Table 13).

Comparisons at time block five revealed that drivers who engaged in a simple verbal task exhibited fewer steering deviations than drivers who did not have a verbal task, $t(28) = 2.56$, $p = .012$, $d = 1.297$. There was no difference in steering deviation between no verbal and complex verbal task, $t(28) = 1.76$, $p = .133$, d = 0.659, and simple verbal and complex verbal tasks, $t(28) = -1.91$, $p = .133$, $d = -0.712$ (see Table 13).

Congested drive. All effects for this measure were non-significant: time, $F(1, 42) = .002$, $p = .966$, $\eta^2_p = .0$; verbal task load, $F(2, 42) = .89$, $p = .42$, $\eta^2_p = .04$; and time by verbal task load interaction, $F(2, 42) = 1.54$, $p = .227$, $\eta^2_p = .068$.

Effect of time on each level of verbal task load. Planned analysis indicated a reduction in steering deviation between time blocks four and five for drivers who had a complex verbal load, $t(14) = 3.44$, $p = .004$, $d = 0.597$. This reduction in steering deviation was not observed in the no verbal load, $t(14) = 0.13$, $p = .989$, $d = 0.004$, and simple verbal load groups, $t(14) = -0.79$, $p = .439$, $d = -0.255$ (see Table 14).

Table 13[a]
Root mean square error (RMSE) of steering between (in degrees) verbal task loads at time blocks four and five under monotonous conditions.

	Time[b] Block 4	Block 5
Verbal load	RMSE - Steering	
No verbal load	1.28 (.66)	1.26 (.63)
Simple verbal load	0.85 (.41)	0.68 (.27)
Complex verbal load	1.27 (.55)	0.91 (.41)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Table 14[a]
Root mean square error (RMSE) of steering (in degrees) between verbal task loads at time blocks four and five under congested conditions.

	Time[b] Block 4	Block 5
Verbal load	RMSE - Steering	
No verbal load	1.78 (.87)	1.78 (.66)
Simple verbal load	1.51 (.85)	1.88 (1.99)
Complex verbal load	1.56 (.72)	1.18 (.56)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Effect of verbal task load at time block five. Planned analysis revealed that drivers who engaged in a complex verbal task exhibited smaller steering deviations than drivers who did not have a verbal load, $t(28) = 2.66$, $p = .038$, $d = 0.975$. There was no difference in deviation between no verbal and simple verbal loads, $t(28) = -0.19$, $p = .853$, $d = -0.077$, and simple verbal and complex verbal loads, $t(28) = 1.31$, $p = .419$, $d = 0.547$ (see Table 14).

Steering Deflections (> 10°)

Monotonous drive. All effects for this measure were non-significant: time, $F(1, 42) = .0$, $p = .1$, $\eta^2_p = .0$; verbal task load, $F(2, 42) = 1.70$, $p = .194$, $\eta^2_p = .075$; and time by verbal task load interaction, $F(2, 42) = 1.61$, $p = 212$, $\eta^2_p = .071$.

Effect of time on each level of verbal task load. Planned analysis did not reveal significant changes in steering deflection between time blocks four and five for the no verbal, $t(14) = -0.87$, $p = .397$, $d = -0.168$; simple verbal, $t(14) = 1.29$, $p = .217$, $d = 0.668$; and complex verbal loads, $t(14) = 1.66$, $p = .120$, $d = 0.581$ (see Table 15).

Effect of verbal task load at time block five. Planned analysis did not reveal significant differences in deflections between no verbal and simple verbal loads, $t(28) = 1.38$, $p = .189$, $d = 0.712$; no verbal and complex verbal loads, $t(28) = 1.25$, $p = .220$, $d = 0.599$; and simple verbal and complex verbal loads, $t(28) = -1.44$, $p = .173$, $d = -0.741$ (see Table 15).

Table 15[a]
Steering deflections > 10° under monotonous conditions.

	Time[b] Block 4	Block 5
Verbal load	Steering deflections	
No verbal load	2.47 (5.21)	3.80 (10.67)
Simple verbal load	0.27 (.79)	0.00 (.00)
Complex verbal load	1.40 (2.77)	0.33 (.89)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Congested drive. All effects for this measure were non-significant: time, $F(1, 42) = .75$, $p = .391$, $\eta^2_p = .018$; verbal task load, $F(2, 42) = .51$, $p = .61$, $\eta^2_p = .020$; and time by verbal task load interaction, $F(2, 42) = .87$, $p = .43$, $\eta^2_p = .040$

Effect of time on each level of verbal task load. Planned analysis did not reveal significant changes in steering deflections between time blocks four and five for the no verbal, $t(14) = -0.88$, $p = .389$, $d = -0.277$; simple verbal, $t(14) = -0.89$, $p = .391$, $d = -0.309$, and complex verbal load groups, $t(14) = 1.97$, $p = .069$, $d = 0.413$ (see Table 16).

Effect of verbal task load at time block five. Planned analysis did not reveal significant differences in deflections between no verbal and simple verbal loads, $t(28) = -0.47$, $p = .644$, $d = -0.198$; no verbal and complex verbal loads, $t(28) = 1.86$, $p = .078$, $d = 0.739$, and simple verbal and complex verbal loads, $t(28) = 1.01$, $p = .331$, $d = 0.472$ (see Table 16).

Table 16[a]
Steering deflections > 10˚ under congested conditions.

	Time[b] Block 4	Block 5
Verbal load	Steering deflections	
No verbal load	2.50 (3.56)	3.73 (4.62)
Simple verbal load	2.47 (4.75)	5.93 (17.62)
Complex verbal load	2.33 (2.97)	1.33 (1.88)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Lane Infraction Count
Monotonous drive. There was no main effect of time, $F(1, 42) = 2.77$, $p = .103$, $\eta^2_p = .062$, neither was there a significant time by verbal task load interaction, $F(2, 42) = 2.48$, $p = .096$, $\eta^2_p = .106$. However, there was a significant main effect of verbal task load, $F(2, 42) = 6.08$, $p = .005$, $\eta^2_p = .225$

Effect of time on each level of verbal task load. Planned analysis did not reveal significant changes in lane infractions between time blocks four and five for the no verbal load group, $t(14) = - 0.53$, $p = .604$, d = -.0.09. There was a marginal reduction of infractions in the complex verbal load group, $t(14) = 2.08$, $p = .057$, $d = 0.764$. Drivers in the simple verbal load group exhibited a significant reduction in infractions, $t(14) = 2.87$, $p = .012$, $d = 0.753$ (see Table 17).

Table 17[a]
Number of lane infractions at time blocks four and five under monotonous conditions.

	Time[b] Block 4	Block 5
Verbal load	Infractions	
No verbal load	7.33 (7.55)	8.13 (10.07)
Simple verbal load	1.87 (2.07)	0.60 (1.29)
Complex verbal load	4.33 (6.24)	1.13 (2.13)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Effect of verbal task load at time block five. Planned analysis revealed that drivers in the simple verbal load group made fewer infractions than drivers in the no verbal load group, $t(28) = 2.87$, $p = .036$, $d = 1.326$. Similarly, drivers in the complex verbal load group also exhibited fewer infractions than drivers in the no verbal load group, $t(28) = 2.63$, $p = .037$, $d = 1.147$. There was no difference in lane infractions between drivers in the simple and complex verbal load groups, $t(28) = - 0.83$, $p = .415$, $d = - 0.311$ (see Table 17)

Congested Drive. All effects for this measure were non-significant: time, $F(1, 42) = .23$, $p = .631$; $\eta^2_p = .006$, verbal task load. $F(2, 42) = .39$, $p = .674$, $\eta^2_p = .019$; and time by verbal task load interaction, $F(2, 42) = .85$, $p = .435$, $\eta^2_p = .039$.

Effect of time on each level of verbal task load. Planned analysis did not reveal significant changes in lane infractions between time blocks four and five for no verbal load, $t(14) = - 1.00$, $p = .334$, $d = -0.212$; simple verbal load , $t(14) = - 0.71$, $p = .492$, $d = - 0.262$; and complex verbal load groups, $t(14) = 1.84$, $p = .088$, $d = 0.516$ (see Table 18).

Effect of verbal task load at time block five. Planned analysis did not reveal significant differences in infractions between no verbal and simple verbal load, $t(28) = -0.49$, $p = .627$, $d = 0.212$; no verbal and complex verbal loads, $t(28) = 1.83$, $p = .082$, $d = 0.712$; and simple verbal and complex verbal loads, $t(28) = 0.97$, $p = .347$, $d = 0.449$ (see Table 18).

Table 18[a]
Number of lane infractions at time block four and five under congested conditions.

	Time[b] Block 4	Block 5
Verbal load	Infractions	
No verbal load	0.67 (1.23)	0.93 (1.28)
Simple verbal load	0.73 (1.49)	1.67 (5.64)
Complex verbal load	0.80 (1.47)	0.27 (0.69)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Lane Infraction Duration

Monotonous Drive. There was no main effect of time, $F(1, 42) = 3.48$, $p = .069$, $\eta^2_p = .076$, neither was there a significant time by verbal task load interaction, $F(2, 42) = 1.12$, $p = .335$, $\eta^2_p = .051$. However, there was a significant main effect of verbal task load, $F(2, 42) = 7.19$, $p = .002$, $\eta^2_p = .254$.

Effect of time on each level of verbal task load. Planned analysis did not reveal significant changes in the duration of lane infractions between time blocks four and five for the no verbal load group, $t(14) = 0.22$, $p = .983$, $d = 0.003$. Drivers in the complex verbal load group did not exhibit a significant change in the duration of infractions, $t(14) = 1.84$, p

= .088, d = 0.615. Drivers who engaged in a simple verbal load exhibited significantly shorter durations in an infraction, $t(14) = 2.66$, $p = .018$, $d = 0.858$ (see Table 19).

Effect of verbal task load at time block five. Planned analysis revealed that drivers in the simple verbal load group spent less time in an infraction than drivers in the no verbal load group, $t(28) = 2.97$, $p = .030$, $d = 1.399$. Similarly, drivers in the complex verbal load group also exhibited shorter durations in an infraction than drivers in the no verbal load group, $t(28) = 2.51$, $p = .046$, $d = 1.049$. There was no difference between drivers in the simple and complex verbal load groups, $t(28) = - 1.21$, $p = .243$, $d = - 0.489$ (see Table 19).

Table 19[a]
Duration of lane infractions in milliseconds (msec) at time block 4 and 5 under monotonous conditions.

	Time[b] Block 4	Block 5
Verbal load	Infraction duration (msec)	
No verbal load	11930 (11692)	11885 (14521)
Simple verbal load	2486 (2692)	700 (1470)
Complex verbal load	5703 (7618)	2078 (4167)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Congested drive. All effects for this measure were non-significant: time, $F(1, 42) = .08$, $p = .774$, $\eta^2_p = .002$; verbal load, $F(2, 42) = .26$, $p = .775$, $\eta^2_p = .012$; and time by verbal load interaction, $F(2, 42) = 1.21$, $p = .308$, $\eta^2_p = .055$.

Effect of time on each level of verbal task load. Planned analysis did not reveal significant changes in lane infractions between time block four and five for the no verbal, $t(14) = -1.34$, $p = .201$, $d = - 0.234$; simple verbal, $t(14) = - 0.62$, $p = .542$, $d = - 0.215$; and the complex verbal load groups, $t(14) = 1.62$, $p = .128$, $d = 0.527$ (see Table 20).

Effect of verbal task load at time block five. Planned analysis did not reveal significant differences in infraction duration between no verbal and simple verbal loads, $t(28) = - 0.91$, $p = .928$., d = - 0.035; no verbal and complex verbal loads, $t(28) = 1.83$, $p = .085$, $d = 0.756$; and simple verbal and complex verbal loads, $t(28) = 1.00$, $p = .325$, $d = 0.456$ (see Table 20).

Table 20[a]
Duration of lane infractions in milliseconds (msec) at time block four and five under congested conditions.

	Time[b] Block 4	Block 5
Verbal load	Infraction duration (msec)	
No verbal load	1242 (2538)	1891 (3010)
Simple verbal load	1146 (2115)	2056 (6338)
Complex verbal load	1470 (3129)	400 (934)

[a] Numbers in parentheses are standard deviations
[b] Each time block is 9 mins in duration

Combined analysis between driving conditions
Unlike the preceding analyses, the following variables were not affected by the systematic differences in speed between driving condition. Unless otherwise stated, the following analyses were conducted using a 2 (Drive load) x 3 (Verbal task load) x 2 (Time) mixed ANOVA.

Attention – Response time to peripheral detection task.
No statistically significant effects were observed on all dependent measures. There was no main effect of time, $F(1, 56) = 0.10$, $p = .748$, $\eta^2_p = .002$; driving load, $F(1, 56) = 0.90$, $p = .346$, $\eta^2_p = .016$; or verbal task load, $F(2, 56) = 0.19$, $p = .823$, $\eta^2_p = .007$. All interactions were non-significant: Driving load x Verbal task load, $F(2, 56) = 1.29$, $p = .284$, $\eta^2_p = .044$; Driving load x Time, $F(1, 56) = 1.41$, $p = .240$, $\eta^2_p = .025$; and Verbal task load x Time, $F(2, 56) = 0.68$, $p = .935$. The three-way interaction was not significant, $F(2, 56) = 0.86$, $p = .429$, $\eta^2_p = .030$.

Planned analysis using a 2 (Driving condition) x 3 (Verbal task load) ANOVA was conducted at time block five to investigate the combined effects of verbal task load and driving load. There was no effect of driving load, $F(1, 68) = 1.67$, $p = .201$, $\eta^2_p = .024$ and no effect of verbal task load, $F(2, 68) = 0.15$, $p = .865$, $\eta^2_p = .004$. The driving load by verbal task load interaction was not significant, $F(2, 68) = 2.45$, $p = .094$, $\eta^2_p = .067$ (See Table 21).

Table 21[a]
Reaction time in milliseconds (msec) to the peripheral detection task (PDT)[b] at time block five.

	Driving load	
	Monotonous	Congested
Verbal Load	Reaction time (msec)	
No Verbal	1411 (1139)	1829 (1012)
Simple Verbal	1982 (1406)	1178 (252)
Complex Verbal	1750 (1023)	1199 (735)

[a] Numbers in parentheses are standard deviations
[b] $N = 74$

Attention – Peripheral Detection Task (PDT)
Chi-square analysis was conducted on time block five to investigate responses to the peripheral detection targets. There was no significant relationship in responses for either driving load, $\chi^2(1, N = 90) = .278$, $p = .793$, Cramer's $V = .056$, or verbal task load, $\chi^2(2, N = 90) = 4.53$, $p = .118$, Cramer's $V = .236$. A second analysis was run to examine the relationship between verbal load and responses controlling for the level of driving condition. While there were no significant relationships in the monotonous driving load, $\chi^2(2, N = 45) = .44$, $p = 1$, Cramer's $V = .082$, there was a significant relationship in responses in the congested driving load, $\chi^2(2, N = 45) = 7.02$, $p = .049$, Cramer's $V = .421$. Follow-up analysis at the level of congested driving load, before correction for multiple comparisons, found a significant relationship in responses between no verbal and simple verbal loads, $\chi^2(1, N = 30) = 6.14$, $p = .035$, Cramer's $V = .452$. Drivers who did not have a verbal load had a 93.3% rate of response. Drivers engaged in a simple verbal load had a response rate of 53.3%. However, after correction, this relationship was no longer significant ($p = .105$). There was no significant relationship in responses between no verbal and complex verbal loads, $\chi^2(1, N = 30) = 0.37$, $p = 1$, Cramer's $V = .111$, and simple verbal and complex verbal loads, $\chi^2(1, N = 30) = 3.97$, $p = .109$, Cramer's $V = .364$.

Attention and Memory – Billboard Recall
Time in this analysis was between the first and last time blocks. There was a significant main effect of time, $F(1, 84) = 3.97$ $p = .050$, $\eta^2_p = .045$, driving load, $F(1, 84) = 10.77$, $p = .002$, $\eta^2_p = .114$, and verbal task load, $F(2, 84) = 5.25$, $p = .007$, $\eta^2_p = .111$. There was a significant time by verbal task load interaction, $F(2, 84) = 10.55$, $p < .001$, $\eta^2_p = .201$, and a significant three-way interaction, $F(2, 84) = 3.35$, $p = .040$, $\eta^2_p = .074$. Simple interactions tests were conducted to investigate the significant three-way interaction.
Effect of verbal task load on time and driving load.
No verbal load. There was a marginal effect of time, $F(1, 28) = 3.91$, $p = .058$, $\eta^2_p = .122$, and driving load, $F(1, 28) = 4.15$, $p = .051$, $\eta^2_p = .129$. There was no significant time by driving load interaction, $F(1, 28) = 0.122$, $p = .730$, $\eta^2_p = .004$ (see Table 22).

Complex verbal load. There was an effect of time, $F(1, 28) = 18.97$, $p < .001$, $\eta^2_p = .404$. Billboard recall declined over time for both groups. However, there was no effect of driving load, $F(1, 28) = 2.09$, $p = .159$, $\eta^2_p = .070$, or a significant time by driving load interaction, $F(1, 28) = 0.65$, $p = .427$, $\eta^2_p = .023$ (see Table 22).

Simple verbal load. Results indicated a main effect of time $F(1, 28) = 5.97$, $p = .021$, $\eta^2_p = .176$, and driving load, $F(1, 28) = 5.11$, $p = .032$, $\eta^2_p = .154$. The time by driving load interaction was significant, $F(2, 28) = 9.86$, $p = .004$, $\eta^2_p = .260$ (see Figure 10). The interaction was investigated by testing the effect of time on each level of driving load, and the effect of driving load at each time block.

Effect of time on driving load. Paired sample t-tests revealed that billboard recall under monotonous driving loads was greater in the first time block than the last time block, $t(14) = 4.29$, $p < .001$, $d = 1.219$. Recall did not differ significantly between the first and last time blocks under congested driving loads, $t(14) = - 0.46$, $p = .653$, $d = - 0.145$ (see Table 22).

Effect of driving load on time. There was a significant difference in billboard recall between monotonous and congested loads at the first time block, $F(1, 28) = 12.06$, $p = .002$, $\eta^2_p = .301$. There was no difference in the proportion of billboards recalled in the final block, $F(1, 28) = 0$, $p = 1$, $\eta^2_p = .0$ (see Table 22).

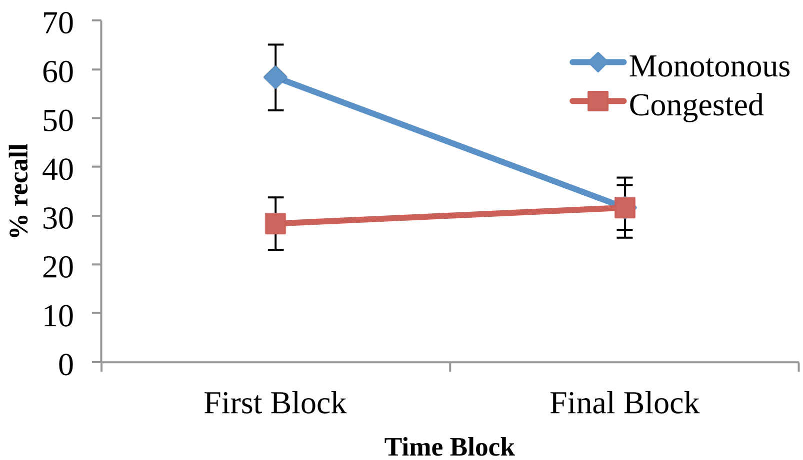


Figure 10[a]. Time by drive load interaction at the level of simple verbal load for billboard recall. Bars represent standard error of the mean.

Effect of driving load on time and verbal task load.

Monotonous drive. There was a marginal effect of verbal task load, $F(2, 42) = 2.85$ $p = .069$, $\eta^2_p = .119$, and a a significant time by verbal task load interaction, $F(1, 42) = 6.56$ $p = .003$, $\eta^2_p = .238$ (see Figure 11). There effect of time was not significant, $F(1, 42) = 3.32$, $p = .075$, $\eta^2_p = .073$. The interaction was investigated by testing the effect of time on each level of verbal load, and the effect of verbal load at each time block.

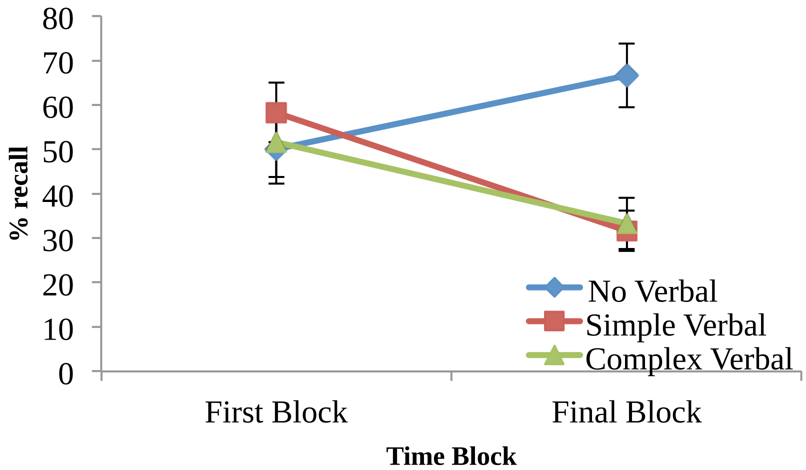


Figure 11[a]. Time by verbal task load interaction at the level of monotonous conditions for billboard recall. Bars represent standard error of the mean.

Effect of time on verbal task load. Paired sample t-tests revealed that billboard recall was better in the first time block than final time block for drivers who engaged in a simple verbal load, $t(14) = 4.29$, $p < .001$, $d = 1.219$. Drivers who engaged in a complex verbal load had a marginally higher proportion of recall in the first time over the final time block, $t(14) = 1.98$, $p = .068$, $d = 0.691$. (see Table 22).

Table 22[a]
Descriptive statistics (%) for simple interactions test for billboard recall.

	Time	
	First Block	Final Block
Driving Load	Billboard recall %	
No Verbal Load		
Monotonous	50.00 (29.88)	66.67 (27.82)
Congested	38.33 (26.50)	50.00 (25.00)
Simple Verbal Load		
Monotonous	58.33 (26.16)	31.67 (17.59)
Congested	28.33 (20.85)	31.67 (24.03)
Complex Verbal Load		
Monotonous	51.67 (30.57)	33.33 (22.49)
Congested	45.00 (27.06)	18.33 (17.59)

[a] Numbers in parentheses are standard deviations

Effect of verbal task load on time blocks. There was a difference in the proportion of billboards recalled between verbal loads at the final time block, $F(2, 42) = 11.04$, $p < .001$, $\eta^2_p = .345$. Independent samples t-test at the final time block revealed that drivers in the no verbal load group recalled more billboards than drivers in the simple verbal load, $t(28) = 4.12$, $p < .001$, $d = 1.542$, and complex verbal load groups, $t(28) = 3.61$, $p = .002$, $d = 1.326$. Billboard recall did not differ at the first time block, $F(2, 42) = 0.35$, $p = .708$, $\eta^2_p = .016$ (see Table 22).

Congested drive. There was no effect of time, $F(1, 42) = 0.84$, $p = .364$, $\eta^2_p = .020$, and verbal task load. $F(2, 42) = 2.49$, $p = .095$, $\eta^2_p = .106$. There was a significant time by verbal task load interaction, $F(2, 42) = 7.53$, $p = .002$, $\eta^2_p = .264$ (see Figure 12). The interaction was investigated by testing the effect of time on each level of verbal task load, and the effect of verbal task load at each time block.

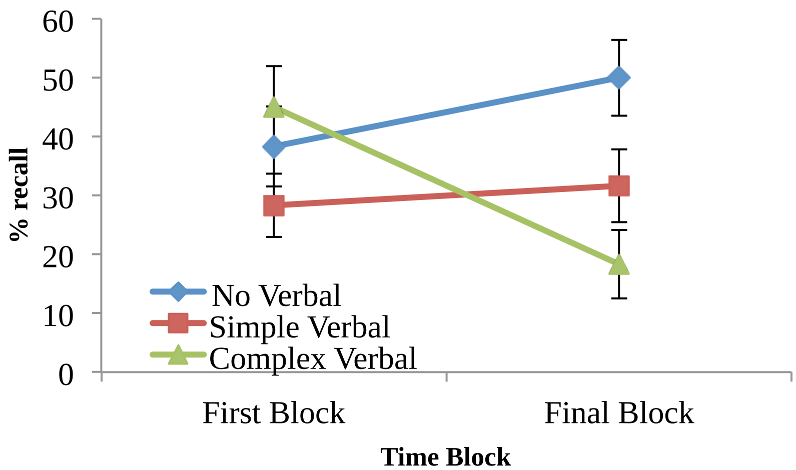


Figure 12[a]. Time by verbal task load interaction at the level of congested conditions for billboard recall. Bars represent standard error of the mean.

Effect of time on verbal task load. Paired sample t-tests revealed that drivers who engaged in a complex verbal load recalled fewer billboards, $t(14) = 5.87$, $p < .001$, $d = 1.195$, in the final time block than the first time block. There were no significant changes in billboard recall between the first and final time blocks when drivers did not have a verbal load, $t(14) = -1.24$, $p = .235$, $d = -0.453$, or engaged in a simple verbal load, $t(14) = -0.46$, $p = .653$, $d = 0.315$ (see Table 22).

Effect of verbal task load on time blocks. There was a difference in the proportion of billboards recalled between verbal loads at the final time block, $F(2, 42) = 7.52$, $p = .002$, $\eta^2_p = .264$. Independent samples t-test at the final time block revealed that drivers in the no verbal load group recalled more billboards than drivers in the complex verbal load group, $t(28) = 4.01$, $p = .001$, $d = 1.487$. Billboard recall did not differ between the simple verbal load group and the no verbal load group, $t(28) = 2.05$, $p = .100$, $d = 0.748$. In addition, there was no difference between simple and complex verbal load groups, $t(28) = 1.73$, $p = .100$, $d = 0.641$. There was no difference in the proportion of billboard recall in the first time block, $F(2, 42) = 1.69$, $p = .196$, $\eta^2_p = .075$ (see Table 22).

Effect of time on driving load and verbal task load.

Time block one. There was a significant main effect of driving load, $F(1, 84) = 7.99$, $p = .006$, $\eta^2_p = .087$. There was no effect of verbal task load, $F(2, 84) = 0.29$, $p = .745$, $\eta^2_p = .007$. The interaction between interaction between verbal task load and drive load was not significant, $F(2, 84) = 1.55$, $p = .218$, $\eta^2_p = .036$.

Time block two. There was a significant main effect of driving load, $F(1, 84) = 4.85$, $p = .030$, $\eta^2_p = .055$, and verbal task load, $F(2, 84) = 17.43$, $p < .001$, $\eta^2_p = .293$. The interaction between interaction between verbal task load and drive load was not significant, $F(2, 84) = 1.23$, $p = .300$, $\eta^2_p = .028$.

Task engagement (TE) and NASA-TLX scores

Only the effect of time was significant, $F(1, 84) = 88.83$ $p < .001$, $\eta^2_p = .514$. Post-drive TE scores ($M = 77.14$, $SD = 16.83$) were significantly lower than pre-drive scores ($M = 90.32$, $SD = 10.09$). There was no effect of driving load, $F(1, 84) = 0.50$, $p = .480$, $\eta^2_p = .006$, or verbal task load, $F(2, 84) = 2.61$, $p = .080$, $\eta^2_p = .058$. All interactions were non-significant: Driving load x Verbal task load, $F(2, 84) = 2.43$, $p = .094$, $\eta^2_p = .055$; Driving load x Time, $F(1, 84) = 0.09$, $p = .924$, $\eta^2_p = .000$; and Verbal task load x Time, $F(2, 84) = 2.48$, $p = .090$, $\eta^2_p = .056$. The three-way interaction was not significant, $F(2, 84) = 2.59$, $p = .083$, $\eta^2_p = .057$.

Planned analysis at the final time block using a 2 (Driving load) x 3 (Verbal task load) mixed factorial ANOVA on post drive task engagement scores revealed a main effect of verbal task load, $F(2, 84) = 3.31$, $p = .041$, $\eta^2_p = .073$. There was a significant driving load by verbal task load interaction, $F(2, 84) = 3.22$, $p = .045$, $\eta^2_p = .071$ (see Figure 13). There was no effect of driving load, $F(1, 84) = 0.23$, $p = .634$, $\eta^2_p = .003$. The interaction was investigated by testing the effect of driving load at each level of verbal task load and the effect of verbal task load at each driving load.

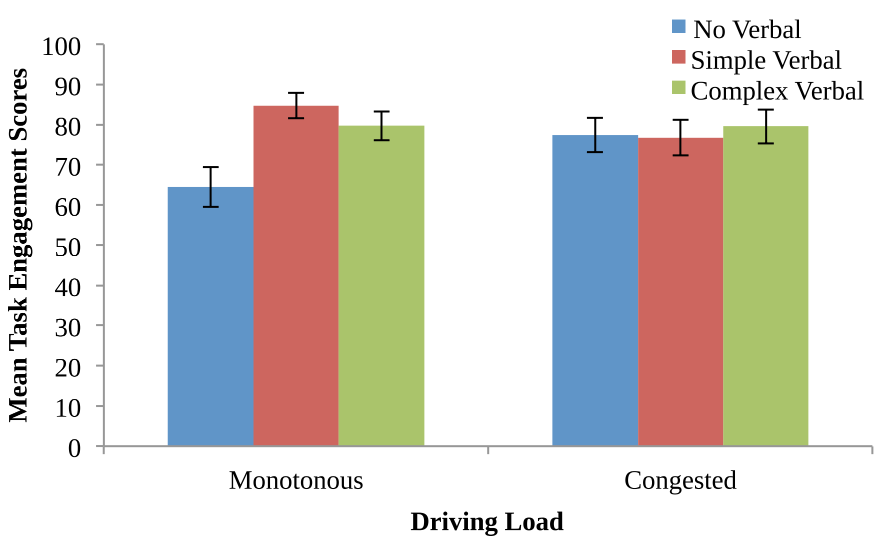


Figure 13[a]. Task engagement scores for the interaction of driving load and verbal task load at time block five. Bars represent standard error of the mean.

Effect of driving load on each level of verbal task load. Drivers who drove in congestion and did not engage in a verbal task reported higher task engagement (TE) scores, than drivers in the monotonous condition who similarly did not engage in a verbal task, $F(1, 84) = 4.84$, $p = .031$, $\eta^2_p = .054$, There was no difference in TE scores between driving load for drivers who engaged in the simple, $F(1, 84) = 1.82$, $p = .181$, $\eta^2_p = .021$, or complex verbal task, $F(1, 84) = 0.001$, $p = .982$, $\eta^2_p = .000$ (see Table 23).

Table 23[a]
Task engagement (TE) scores for driving load and verbal task load at time block five.

	Driving Load Monotonous	Congested
Verbal Task Load	Task engagement scores	
No Verbal	64.53 (19.11)	77.47 (16.71)
Simple Verbal	84.73 (12.09)	76.80 (17.16)
Complex Verbal	79.73 (14.09)	79.60 (16.47)

[a] Numbers in parentheses are standard deviations

Effect of verbal task load on each level of driving load. Analysis of the interaction indicated a significant difference in task engagement (TE) scores between drivers in the monotonous condition, $F(2, 84) = 6.41$, $p = .003$, $\eta^2_p = .132$. This difference was not significant under congested driving conditions, $F(2, 84) = 0.12$, $p = .884$, $\eta^2_p = .003$. Independent samples t-test conducted at the level of monotonous driving conditions revealed that drivers in the no verbal load group reported significantly lower TE scores than drivers in the simple, $t(28) = - 3.46$, $p = .005$, $d = - 1.551$, and complex verbal load groups, $t(28) = - 2.48$, $p = .038$, $d = - 0.916$. Performance did not differ between simple and complex verbal load groups, $t(28) = 1.04$, $p = .306$, $d = - 0.688$ (See Table 10).
There was no significant difference in the overall NASA-TLX scores across the six possible conditions, $F(5, 84) = 0.69$, $p = .636$, $\eta^2_p = .039$, ($M = 4.70$, SD = 1.35).

Control group task engagement and NASA-TLX scores

Analysis of task engagement (TE) scores for the control group did not reveal any significant effects. There was no significant change between pre-task ($M = 87.22$, $SD = 15.09$) and post-task ($M = 85.44$, $SD = 16.56$) TE scores for participants in the control group, $F(1, 30) = 0.76$, $p = .389$, $\eta^2_p = .025$. Scores between the simple verbal load ($M = 87.63$, $SD = 16.14$) and complex verbal load ($M = 85.03$, $SD = 15.85$) levels of the control group were not statistically significant, $F(1, 30) = 0.199$, $p = .660$, $\eta^2_p = .007$. Additionally, there was also no significant difference in overall NASA-TLX scores between the two levels of the control group, $t(30) = 0.59$, $p = .581$, $d = 0.196$.

Comparisons between controls and treatment on task engagement

Task engagement (TE) scores from the control group were compared against 32 randomly selected participants from the experimental group. Results indicated a significant difference between pre and post-task scores, $F(1, 62) = 34{,}64$, $p < .001$, $\eta^2_p = .358$, and a significant time by group interaction , $F(1, 84) = 23.16$, $p < .001$, $\eta^2_p = .272$. The interaction was investigated by testing the effect of group (control or experimental) at pre and post task TE scores, and the effect pre and post task TE scores on the level of group.

Effect of group (control or experimental) at pre and post task scores. Participants in experimental group reported a significant decrease in task engagement (TE) scores between pre-task ($M = 90.56$, $SD = 10.03$) and post task (M = 72.81, $SD = 17.85$), $t(31) = 6.72$, $p < .001$, $d = 1.273$. The control group did not show a significant change between pre ($M = 87.22$, $SD = 15.09$) and post tasks ($M = 85.44$, $SD = 16.56$) TE scores, $t(31) = 0.89$, $p = .383$, $d = 0.112$.

Effect of pre and post task scores for groups (control or experimental). At post-task, participants in the experimental group (M = 72.81, $SD = 17.85$) reported lower task engagement (TE) scores, $F(1, 62) = 8.61$, $p < .001$, $\eta^2_p = .122$, than participants from the control group ($M = 85.44$, $SD = 16.56$). Pre-task TE scores did not differ, $F(1, 62) = 1.09$, $p = .301$, $\eta^2_p = .017$, between experimental ($M = 90.56$, $SD = 10.03$) and control group ($M = 87.22$, $SD = 15.09$).

There was no significant difference in NASA-TLX scores, $t(62) = 1.63$, $p = .107$, $d = 0.415$, between experimental ($M = 4.38$, $SD = 1.79$) and control groups ($M = 3.72$, $SD = 1.39$).

Section 5: Conclusions

Using the metaphor of cognitive resources, this project sought to address two overarching questions with respect to vigilance and driving performance: (a) how does perceptual load influence driving performance? (b) How does a concurrent verbal task load influence driving performance particularly when vigilance is low?

Firstly, there is evidence suggesting that higher perceptual loads can attenuate performance decrements by increasing engagement to the environment. Drivers in congested conditions may feel the need to drive safely by having to avoid other vehicles on the roadway. This results in greater awareness of their immediate surroundings and increased focus on the drive. Such behavior is likely achieved by allocating the appropriate amount of resources to the driving task. In contrast, drivers in the monotonous conditions may have a false sense of safety as there is little traffic to worry about. As such, allocation of resources to the driving task is reduced. This leads to a reduction of engagement to the task of driving and an increase in task-unrelated behaviors and thoughts, thus resulting in impaired performance.

Secondly, results indicate improved driving performance when a concurrent verbal task is introduced to mitigate performance decline when vigilance is low. However, this improvement is not without a cost. The benefits of a concurrent verbal task are clearly seen in lane keeping performance. Drivers who engage in a concurrent verbal task, regardless of driving condition, are better at keeping their vehicle in a straight line and are less likely to weave about within the lane. In terms of resources, this result is opposite of what the depletion model would predict. Rather than further reducing performance, the addition of a verbal load improved it. However, the cost to this improvement is a reduction of attention to objects in the periphery. The allocation of resources to visual attention under both cognitive and perceptual load when vigilance is low is unclear at this point and future studies as mentioned in the previous section should prove fruitful.

The results of this dissertation also suggest that both redirection and depletion models of resources are tenable in describing and predicting changes to performance. Performance decrements continue to occur regardless of load, suggesting a depletion of resources. However, the attenuation of performance declines along with performance boosts in the presence of increased perceptual or cognitive load respectively, suggests that load can reduce the redirection of resources to task unrelated processes.

In closing, the results of the behavioral portion of this project present a tempting proposition for drivers who find reasons to engage in a cell phone conversation while driving. While the driver is potentially more vigilant to the roadway, the chances of missing objects in the periphery also increase. Thus given the costs, erring on the side of caution is the safe and smarter choice. The best thing a driver can do when vigilance is low is simply to pull over.

Goal 2: Build techniques based on dynamic systems modeling that are more sensitive than traditional approaches to analyze performance and changes in performance accompanying declines in sustained attention.

Goal 3: Using the dynamic models we will develop a preliminary understanding how sustained attention differs across individuals as well as within an individual over time.

Analyses promised in the proposal:

Section 1: Within person factor analyses

Quick summary: Within person factor analyses proved too unstable to make this approach worth pursuing. Despite consideration of multiple variations of the analyses, the majority of individual models did not converge.

Section 2: Use of momentary derivative estimates to evaluate driver performance

Quick summary: Further considered whether momentary derivative estimates could act as a proxy for other measures typically used such as lane variance. Reasonable correlations between lane variance and momentary derivatives within individual level data.

Section 3: Multilevel models of momentary derivative estimates

Quick summary: A stepwise algorithm was used to examine the relationship of momentary derivative estimates with predictors such as time, drive condition, verbal task, etc. Some interesting plots produced.

Additional analyses conducted

Section 4a: Fractal Analysis

Quick summary: Same idea as Section 2, but from a very different perspective.

Section 4b: Multilevel models of fractal estimates

Quick summary: Parallels section 3. Some interesting results

Section 5a: Relating reaction time task to momentary derivative estimates

Quick summary: A relationship on a moment-to-moment basis could not be established between the reaction times on the PDT and the momentary derivative estimates.

Section 5b: Relating reaction time task to fractal estimates

Quick summary: A relationship on a moment-to-moment basis could not be established between the reaction times on the PDT and the fractal estimates.

Section 1: Within person factor analyses

From Grant: "Within person exploratory and confirmatory factor analyses (also called P-technique) will be used model simulator data to extract a measure of *driving performance*. The reasons for doing so are two fold: 1) these analyses will determine which variables best represent driving performance and 2) these analyses will determine whether the same weighting of variables can be used across different individuals or whether different indicators may be more indicative of driving performance for differing individuals. Unlike common factor analyses that examine between-individual variability, P-technique analyses use repeated measures from a single individual to examine within-individual variability[70, 71] and allows researchers to consider between-person variability in within-individual relationships[72]. **These results will allow one to gauge whether the best indicators of *driving performance* differ across individuals, and whether it is necessary to build a model where the indicators adapt specifically to each individual.**"

Summary of Analyses

Data Preparation & Analysis Notes

Results

Summary for Grant

The tenability of a latent *driving performance* factor was examined using within-person factor analyses (P-technique). Analyses were conducted on participants who did not record a crash during the simulation; this was done to avoid disruptions in the time series and to include those who were most compliant with simulation instructions. Participants fitting this criterion were taken from all three non-verbal conditions in the congested simulations. Models were first fit to interval variance estimates (second moment; 2 second increments or 60 Hz) using lateral velocity, steering angle, and longitudinal velocity as indicators of *driving performance*. Factor scores from this *driving performance* factor were correlated with lane deviation to determine whether the factor shared more variance with lane deviation as opposed to the individual indicators. Although this correlation was in many cases higher as opposed the correlation between lane deviation and the individual indicators, the model did not converge for the majority of the sample. A group model was next fit in which participants' series were stacked to form a single elongated series. A sampling window of 18 minutes was estimated over the length of this series 200 times in overlapping increments. Convergence improved for a *driving performance* factor defined by lateral velocity, steering angle, and lane deviation; however, significant amount of the windows still did not converge, especially in the simple and complex tasks conditions. Using longitudinal velocity in place of lane deviation resulted in substantially less converged solutions. In windows that did converge, lateral velocity generally had the highest standardized loading compared to the other indicators. Further, factor scores from the converged solutions showed a theoretical pattern of increasing variance over time, lending support to the validity of the factor. Therefore, although measurement of a latent *driving performance* factor was not tenable with these data, we did find support for using lateral velocity and to a lesser extent steering angle and lane deviation as indicators of participants' driving performance.

Data Cleaning & Variables

Code files: `clean_data.R`

Summary: Data from participants in the monotonous condition were excluded for all of the following analyses. Furthermore, we excluded persons who were involved in at least one collision during the simulation. Table 1 contains information on the number of collisions, speed exceedances, and off-road excursions for all participants in the congested condition. Variables used in the following analyses were: Lateral velocity (ft/sec), steering angle (degrees), longitudinal velocity (ft/sec), longitudinal acceleration (ft/sec^2), and lane deviation (ft). We were not concerned with the direction of deviations for these variables (e.g., leftward vs. rightward angle on steering wheel), and thus each variable was transformed as follows:

Variables were calculated as follows:

Lateral Velocity = zscores(| Lateral Velocity |)
= z-scores of the absolute value of Lateral Velocity

Steering Angle = zscores(| Steering Angle |)
= z-scores of the absolute value of Steering Angle

Longitudinal Velocity = zscores(| Longitudinal Velocity |)
= z-scores of the absolute value of Longitudinal Velocity

Longitudinal Acceleration = zscores(| Longitudinal Acceleration |)
= z-scores of the absolute value of Longitudinal Velocity

Lane Deviation = zscores(|min{(Lane Position) – max(Lane Position), (Lane Position) – min(Lane Position)}|)
= z-scores of the absolute value of the deviation from the center of the car to the edges of the road, whichever is smaller.

z-scores were used because of the disparity in metric (e.g., M = 1.53, .24, 3.26 for steering angle, lateral velocity, lane deviation, M = 40.60 for longitudinal velocity). The R code `clean_data.R` is used to read in the raw data files (for the congested condition), transform the variables, combine all participants' data into a list (crash and non-crash participants), and export the R workspace containing only the list, so that the data can be sourced into other R programs.

Table 1. *Number of Crashes, Speed Exceedances, and Off-road Excursions for Congested Condition*

Participant	Non-Verbal			Simple Verbal			Complex Verbal		
	Crashes	Speed	Off-road	Crashes	Speed	Off-road	Crashes	Speed	Off-road
1	0	0	20	0	5	2	2	3	0
2	0	0	0	0	0	0	6	2	16
3	0	0	2	2	2	1	1	0	0
4	4	5	1	0	12	18	0	0	0
5	0	0	9	0	0	0	0	2	0
6	1	6	8	0	0	9	0	0	0
7	0	0	0	2	3	0	1	0	2
8	1	1	0	0	0	0	0	1	0
9	0	0	0	1	0	0	0	0	0
10	0	0	0	1	1	0	1	3	0
11	0	0	3	2	0	0	0	0	0
12	0	0	1	0	0	0	0	0	0
13	2	21	13	0	0	1	0	3	6
14	0	1	0	0	0	0	0	3	1
15	1	3	0	0	3	0	1	0	6
16	1	0	2	0	11	8	--	--	--
17	1	38	1	0	2	3	--	--	--

Individual P-technique Models for Variance Estimates

Code files: `second_moment_ind.R, congested_original.RData,`

Summary: The hypothesized driving performance factor was evaluated by fitting within-person factor analyses to individuals who did not crash in the congested condition. Because lane deviation is often used as a measure of driving performance, we considered this our "gold standard" and determined the extent to which the driving performance factor correlated more strongly with lane deviation as opposed to the individual indicators themselves. As shown in the tables below, the factor scores generally correlated more strongly with lane deviation as opposed to the individual indicators, although in some cases the difference in some was negligible. However, convergence was poor regardless of whether longitudinal acceleration or longitudinal velocity was used.

Table 5. *Correlations when Longitudinal Velocity was the Third Indicator*

	Participant	M2			M1		
	FS~LD	LA~ST	LA~LO	ST~LO	LA~LD	ST~LD	LO~LD
nv1	NA	0.56	0.01	-0.02	0.70	0.43	0.00
nv2	0.75	0.52	0.18	0.27	0.54	0.56	0.17
nv3	NA	0.53	-0.02	0.04	0.47	0.33	-0.03
nv5	NA	0.26	-0.01	0.12	0.49	0.15	-0.03
nv7	0.68	0.48	0.04	0.32	0.59	0.37	0.03
nv9	NA	0.41	-0.03	0.08	0.53	0.31	-0.02
nv10	0.60	0.71	0.10	0.16	0.53	0.48	0.00

nv11	0.59	0.82	0.13	0.23	0.58	0.48	0.09
nv12	0.71	0.46	0.06	0.12	0.52	0.44	0.07
nv14	0.84	0.50	0.05	0.04	0.81	0.44	0.04
sv1	0.71	0.32	0.26	0.15	0.66	0.23	0.21
sv2	0.62	0.38	0.04	0.11	0.42	0.35	0.00
sv4	0.81	0.75	0.12	0.04	0.77	0.64	0.13
sv5	NA	0.36	-0.02	0.70	0.59	0.18	-0.03
sv6	NA	0.54	-0.05	0.04	0.53	0.32	-0.03
sv8	0.88	0.60	0.35	0.34	0.74	0.68	0.22
sv12	NA	0.56	0.01	0.09	0.53	0.28	0.03
sv13	NA	0.28	-0.02	0.17	0.64	0.16	0.00
sv14	NA	0.53	0.01	0.10	0.68	0.51	0.01
sv15	NA	0.63	0.01	-0.02	0.63	0.39	-0.01
sv16	0.87	0.48	0.11	0.03	0.69	0.53	0.07
sv17	NA	0.68	0.00	0.03	0.72	0.67	0.00
cv3	0.18	0.03	0.43	0.18	0.08	0.29	0.14
cv4	NA	0.67	-0.01	0.29	0.61	0.41	-0.03
cv6	0.08	0.12	0.03	0.37	0.55	0.04	-0.02
cv7	0.40	0.04	0.96	0.06	0.39	0.47	0.40
cv9	NA	0.60	0.06	0.18	0.85	0.50	0.05
cv11	NA	0.65	0.00	0.07	0.29	0.14	0.03
cv12	NA	0.76	0.00	0.02	0.38	0.52	0.02
cv15	NA	0.10	0.95	0.12	0.14	0.44	0.15

LA = Lateral Velocity; ST = Steering Angle; LO = Longitudinal Velocity; LD = Lane Deviation; FS = Factor Scores; nv = No Verbal Task; sv = Simple Verbal Task; cv = Complex Verbal Task

Table 6. *Correlations when Longitudinal Acceleration was the Third Indicator*

	Participant		M2			M1	
	FS~LD	LA~ST	LA~LO	ST~LO	LA~LD	ST~LD	LO~LD
nv1	NA	0.56	-0.02	-0.04	0.70	0.43	-0.02
nv2	NA	0.52	0.00	0.01	0.54	0.56	0.00
nv3	NA	0.53	-0.02	-0.02	0.47	0.33	-0.02
nv5	NA	0.26	0.00	0.04	0.49	0.15	-0.01
nv7	NA	0.48	-0.03	0.00	0.59	0.37	-0.01
nv9	NA	0.41	-0.04	-0.02	0.53	0.31	-0.03
nv10	0.59	0.71	0.28	0.29	0.53	0.48	0.05
nv11	NA	0.82	0.01	0.02	0.58	0.48	-0.02
nv12	NA	0.46	0.01	0.05	0.52	0.44	0.02
nv14	NA	0.50	-0.01	0.02	0.81	0.44	-0.02
sv1	0.68	0.32	0.07	0.07	0.66	0.23	0.03
sv2	0.62	0.38	0.02	0.16	0.42	0.35	-0.01
sv4	NA	0.75	0.03	-0.01	0.77	0.64	0.07
sv5	NA	0.36	-0.02	0.00	0.59	0.18	-0.03
sv6	NA	0.54	-0.02	0.00	0.53	0.32	-0.04
sv8	0.91	0.60	0.08	0.31	0.74	0.68	0.08

sv12	NA	0.56	-0.02	-0.01	0.53	0.28	0.00
sv13	NA	0.28	-0.04	0.03	0.64	0.16	-0.02
sv14	0.00	0.53	0.00	0.03	0.68	0.51	0.00
sv15	NA	0.63	-0.02	-0.02	0.63	0.39	-0.02
sv16	NA	0.48	0.00	0.04	0.69	0.53	0.02
sv17	NA	0.68	-0.01	0.01	0.72	0.67	-0.02
cv3	0.57	0.03	0.03	0.08	0.08	0.29	0.03
cv4	NA	0.67	-0.01	0.03	0.61	0.41	-0.02
cv6	NA	0.12	0.00	0.07	0.55	0.04	-0.02
cv7	1.96	0.04	0.00	0.08	0.39	0.47	0.05
cv9	0.00	0.60	0.01	0.02	0.85	0.50	0.01
cv11	NA	0.65	-0.02	0.00	0.29	0.14	-0.01
cv12	NA	0.76	0.01	0.02	0.38	0.52	0.02
cv15	0.77	0.10	0.00	0.02	0.14	0.44	0.00

LA = Lateral Velocity; ST = Steering Angle; LO = Longitudinal Velocity; LD = Lane Deviation; FS = Factor Scores; nv = No Verbal Task; sv = Simple Verbal Task; cv = Complex Verbal Task

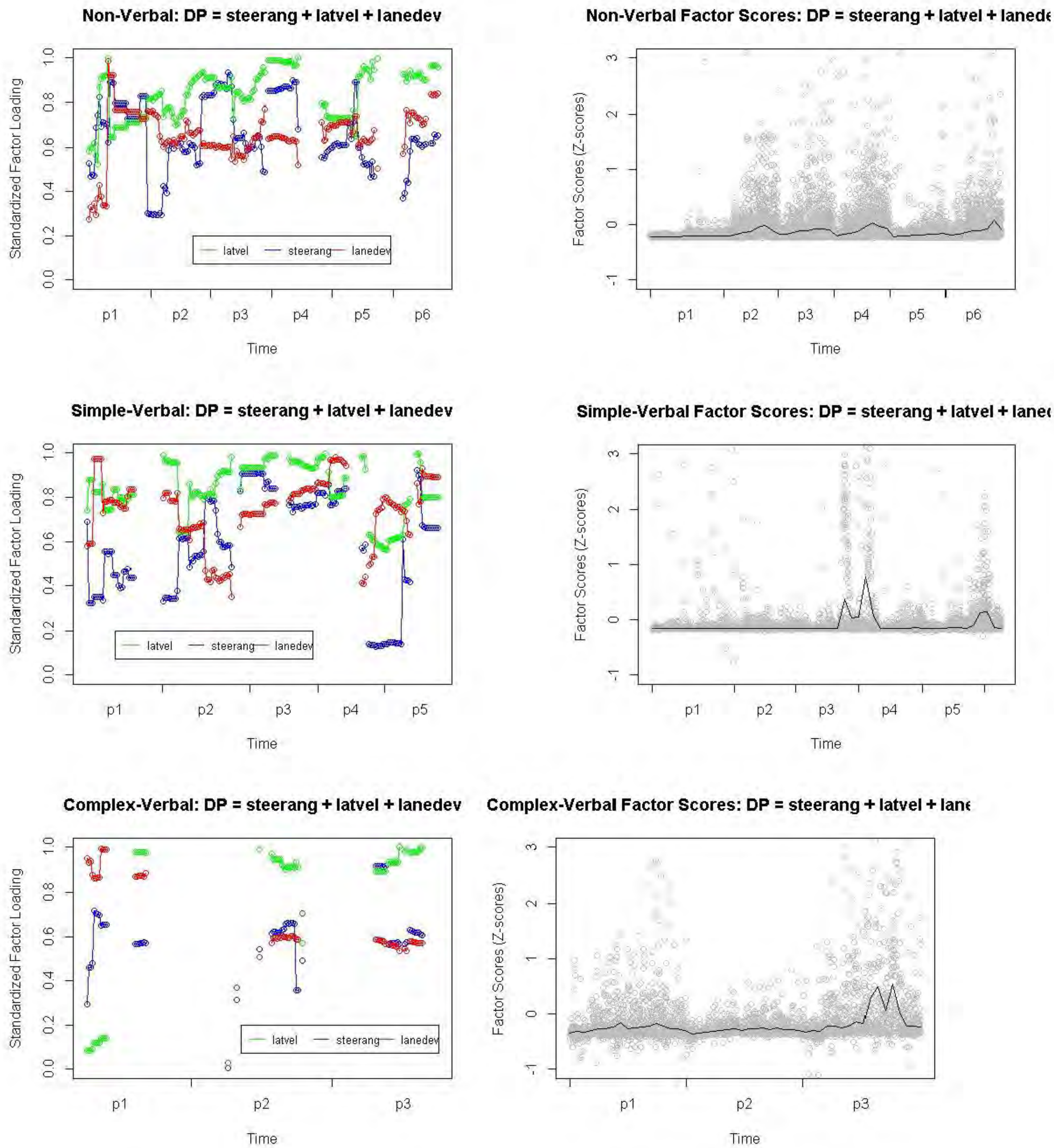
Group P-technique Models for Variance Estimates

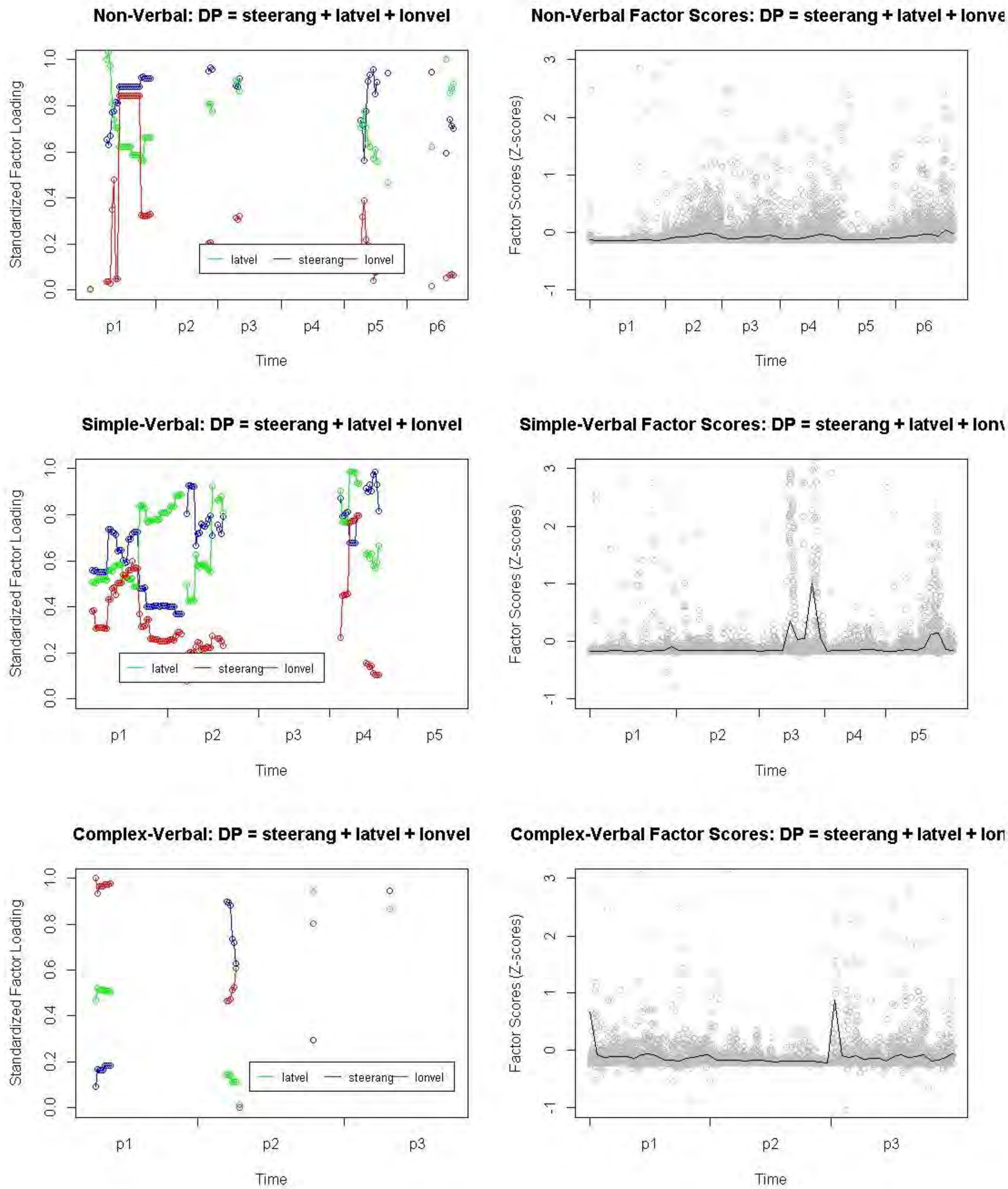
Code files: `group_model_lanedev.R`, `group_model_lonvel.R`, `congested_original.RData`,

Summary: Due to convergence issues with the individual models, several group models were fit to lengthen the time series. These models were fit for the second moments only using 18 minute windows (540 units * 2 second units = 1080 seconds = 18 minutes). For the non-verbal condition, data from participants 2, 7, 10, 11, 12, and 14 were “stacked” in a single vector; For the simple-verbal condition, participants 1, 2, 4, 8, and 16 were stacked; For the complex-verbal condition, participants 3, 6, and 7 were stacked. These participants were chosen because their individual P-technique models converged in earlier conditions. Two sets of models were run: One with lane deviation as the third indicator, and one with longitudinal velocity as the third indicator. Steering angle and lateral velocity were again the other two indicators. The figures below summarize the results from these models. In the left-hand columns, standardized factor loadings are plotted on the y-axis for the length of the time series, which is divided by participant (Note: For the figures, the participants were recoded in ascending order, i.e., non-verbal condition: nv2 = p1, nv7 = p2, and so on). In the right-hand columns, standardized factor scores from each window for the participants were plotted on the y-axis over the length of the group time series. A smoothing function was then used to superimpose a non-parametric density over the scores.

Not surprisingly, the <u>non-verbal condition</u> had the highest rate of converged solutions, purportedly due to the length of the time series. However, there were still “gaps” of non-converged solutions, for example, at the end of the time series for participant nv11 (p4). The relative ordering of loadings was somewhat consistent over the length of the group model time series, with the strongest loadings generally seen for lateral velocity (green); however, the relative ordering of the loadings for steering angle and lane deviation depended on the participant, with higher loadings observed for lane deviation for participants nv7, nv12, and nv14 (red) and higher loadings observed for steering angle for participants nv2, nv10, and nv11 (blue).

With regard to the factor scores, as expected, variation increased toward the end of each time series (i.e., the bumps) likely due to fatigue. In the simple-verbal condition, loadings for lateral velocity and lane deviation were generally highest although the relative ordering was less consistent compared with the non-verbal condition. With regard to factor scores, only participant' sv 4 and sv16 showed marked increases in variance over time. Participant sv8, after a high variance initial period, was able to reduce their volatility in the driving performance factor which increased over time toward the end of the series. No noticeable increases in factor scores were observed for participants sv1 and sv2. The complex-verbal condition had the worst rate of convergence of the three conditions, likely due to the small number of participants (n = 3). For those solutions that did converge, lateral velocity was again the strongest indicator of the driving performance factor, with the exception of the beginning part of the time series for participant cv3. For participants cv6 and cv7, the loadings for steering angle and lane deviation were very similar. The factor score patterns were less discernible than previously, although participant cv7 had a marked increase in driving performance variance towards the end of the series. The next six figures reveal that convergence was extremely poor when longitudinal velocity was used as the third indicator. Interpreting results from this set of models is likely to be misleading.


Non-Verbal: DP = steerang + latvel + lanedev
Standardized Factor Loading
latvel
steerang
lanedev
p1
p2
p3
p4
p5
p6
Time
Non-Verbal Factor Scores: DP = steerang + latvel + lanede
Factor Scores (Z-scores)
Simple-Verbal: DP = steerang + latvel + lanedev
Simple-Verbal Factor Scores: DP = steerang + latvel + lane
Complex-Verbal: DP = steerang + latvel + lanedev
Complex-Verbal Factor Scores: DP = steerang + latvel + lane


Non-Verbal: DP = steerang + latvel + lonvel
Standardized Factor Loading
latvel
steerang
lonvel
p1
p2
p3
p4
p5
p6
Time
Non-Verbal Factor Scores: DP = steerang + latvel + lonve
Factor Scores (Z-scores)
Simple-Verbal: DP = steerang + latvel + lonvel
Simple-Verbal Factor Scores: DP = steerang + latvel + lonv
Complex-Verbal: DP = steerang + latvel + lonvel
Complex-Verbal Factor Scores: DP = steerang + latvel + lon

Section 2: Use of momentary derivative estimates to evaluate driver performance

From Grant: "While derivative estimates are noisy over short time scales, even this simple model suggests that there are patterns in the estimates, including: 1) an increasing number of small corrections over the 90-minute experiment and that the rate of increase diminishes with time, 2) decreases in momentary corrections following an event such as coming to a stop (braking event), and 3) a decreasing number of small corrections following the onset of a dual-task (phone call) late in experiment. **Moment-to-moment estimates of derivatives can be used evaluate whether a driver is making smooth, consistent changes to the vehicle, and these changes appear related to fatigue, attention-grabbing events, and dual-task attention.** This is one of several modeling approaches that will be considered, including models to examine whether drivers are controlling the vehicle in a predictable manner."

Summary of Analyses

For details about this analysis, and its justification, see Deboeck, P. R., Paul A., Chan, M.**, Geldhof, J.**, & Fries, C.** (2011). Using Momentary Derivative Estimates To Gauge Driver Performance. *Advances in Transportation Studies*, 183–192. This article examined how changes in steering angle velocity and absolute lateral velocity related to the standard deviation of lane position. It was demonstrated that estimates of change (first derivative) over short periods of time were very highly correlated with the standard deviation of lane position. The analysis in the paper was expanded for the grant to try to optimize the period over which the change in lateral velocity was estimated.

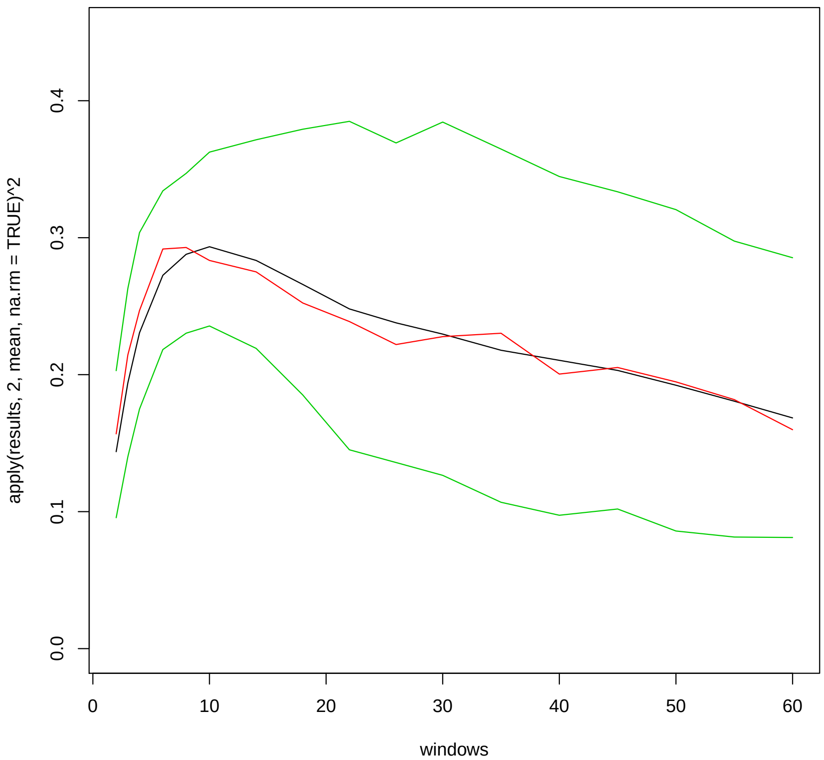
Data Preparation & Analysis Notes

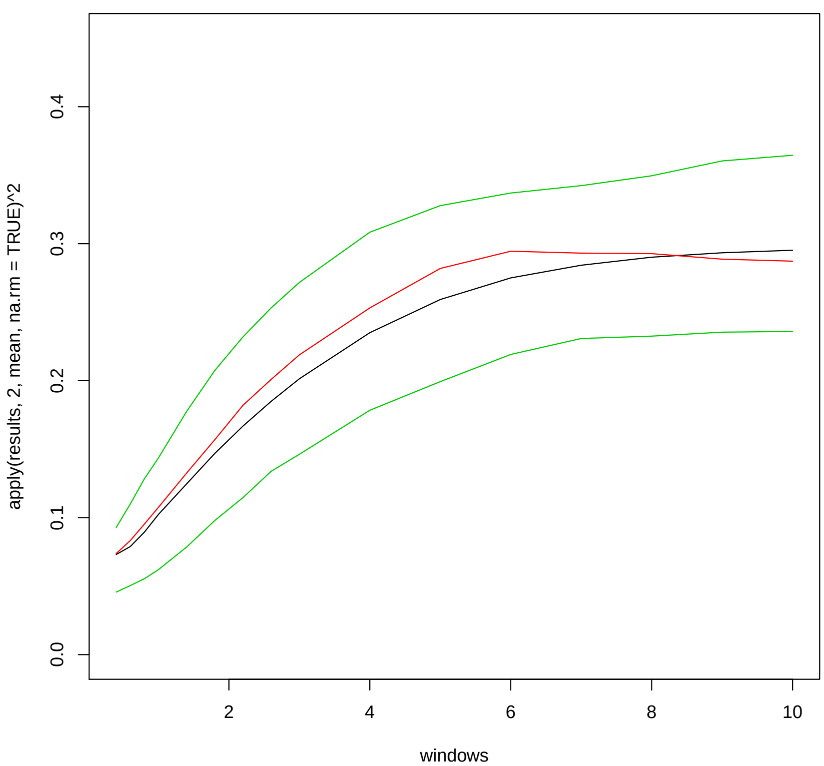
First 5 minutes of data were not analyzed. Data were down sampled to either 2 or 10 measurements per second to make date easier to manage; it was not believed that this would appreciably alter the results. Derivatives were estimated over a span of 2 to 60 seconds, and 0.4 to 10 seconds for each of the down sampled series. Due to the distributions of estimates being very skewed, transformations of the variables were used to clarify the relationship between lateral velocity and lane variance. The square-root of the absolute lateral velocity was compared to the log of lane variance; both estimates were calculated over the same window of time.

The correlation between variables was estimated within individual. The mean across individuals was then examined.

Results

The following plots show the results for the 10 obs/second data (windows 0.4 to 10 seconds) and 2 obs/second data (windows 2 to 60 seconds). The lines represent the mean (black), median (red) and 1st and 3rd quartiles (green) for the correlations across different windows. While the results will depend in part on the non-linear transformations, the results suggest that windows of 6-10 seconds produced the highest average correlation across participants.


apply(results, 2, mean, na.rm = TRUE)^2
0.0
0.1
0.2
0.3
0.4
0
10
20
30
40
50
60
windows


apply(results, 2, mean, na.rm = TRUE)^2
0.0
0.1
0.2
0.3
0.4
2
4
6
8
10
windows

Section 3: Multilevel models of momentary derivative estimates

From Grant: "To examine the effect of experimental conditions on driving performance, and consequently attention, multilevel models will be used. Multilevel models allow one to account for the non-independence of observations in nested data structures; in this case the repeated observations nested within individual would violate the assumptions of ordinary regression[73, 74, 94]. The within person factor analysis will be used to model the ideal combination of driving performance indicators and how they should be weighted for each individual. The dynamical systems model, through the use of a differential equation, will be used to capture the nonlinear, momentary changes in driving performance for each individual. A multilevel component will be added to the differential equation model to estimate the effect of experimental conditions across individuals[75], these experimental conditions include manipulating a variety of driving conditions and vigilance enhancing techniques, in addition to the fatigue expected for many individuals. **The use of methods specifically designed to extract the characteristics of individuals and individual time series, the use of multilevel models to examine interindividual effects while allowing for individual differences, and the variety of experimental conditions will allow this project to determine which factors best mitigate the effects of vigilance declines and when they should be implemented.**"

Summary of Analyses

The analyses also expand upon the work done in Deboeck, P. R., Paul A., Chan, M.**, Geldhof, J.**, & Fries, C.** (2011). Using Momentary Derivative Estimates To Gauge Driver Performance. *Advances in Transportation Studies*, 183–192. A multilevel model was fit to the derivative estimate of the absolute values of the lateral velocity.

Data Preparation & Analysis Notes

Analyses were performed with non-overlapping windows. Attempts were made to fit the models allowing for auto-correlated errors, but the large amount of data led to problems with memory allocation as well as significant convergence problems when the data were downsampled. It would be possible to consider such models, but the data would need to be significantly reduced (e.g., analyzing only every 50th estimate).

Analysis first examined quantile-quantile normal plots to check whether the dependent variable was approximately normally distributed, as there was concern that highly skewed results might bias p-value estimates. Subsequently a series of MLMs were fit with the following characteristics:

The model was run allowing for the following predictors:

- -Drive Condition: congested/monotonous
- -Task Difficulty: No Verbal, Simple Verbal, Complex Verbal
- -Time
- -Time^2: Allowed the effect of time to increase/decrease over time
- -Longitudinal Velocity: Lateral velocity is related to longitudinal velocity, so this can be

considered a covariate
-Intercept Change due to onset of telephone call
-Slope Change due to onset of telephone call

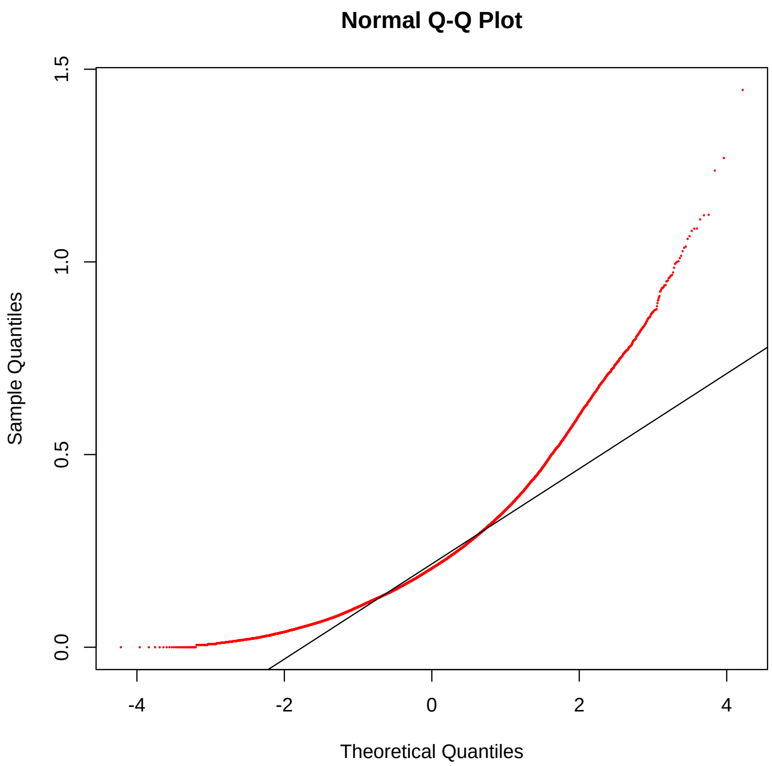
The model was run allowing for the following interactions:
-Drive Condition x Task Difficulty
-Drive Condition x Time
-Drive Condition x Time^2
-Drive Condition x Intercept Change
-Drive Condition x Slope Change
-Task Condition x Intercept Change
-Task Condition x Slope Change
-Drive Condition x Intercept Change x Task Difficulty
-Drive Condition x Slope Change x Task Difficulty

Random Effects were allowed for the following predictors:
-Intercept
-Time
-Time^2
-Intercept Change
-Slope Change

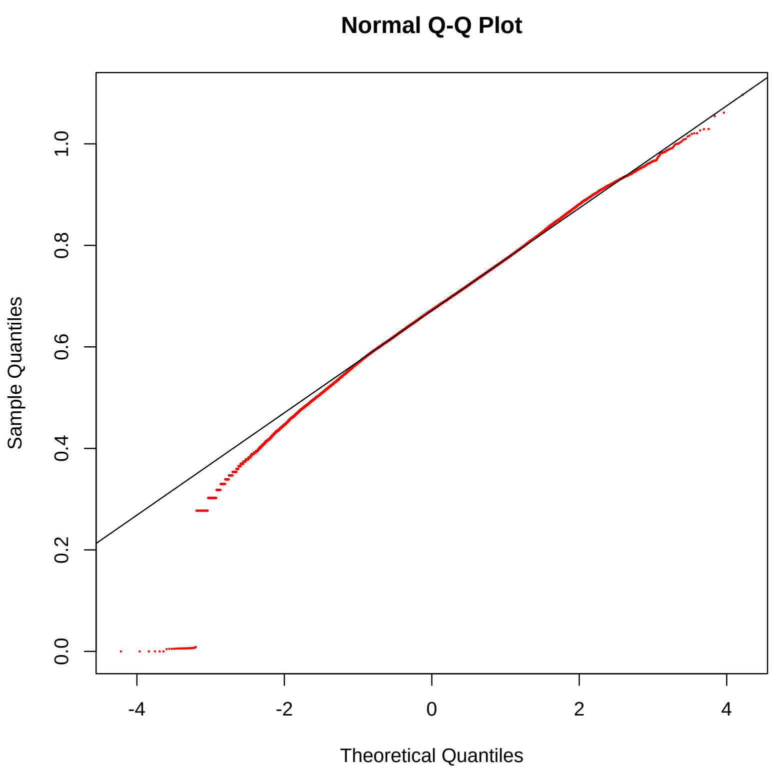
A forward selection algorithm was used to select the ideal set of predictors; this seemed reasonable as initial models suggested that the significant amount of data tends to produce either very significant or clearly non-significant results. The addition of the next variable was based on likelihood improvement. The algorithm used the likelihood ratio test to decide when to stop. For the final model, AIC was used to confirm that there was no evidence suggesting the need for additional predictors. The initial model consisted of an intercept and all of the random effects.

Results

Initial examination of the quantile-quantile normal plots confirmed that the distribution of momentary derivative estimates of absolute lateral velocity were unlikely to have normally distributed errors (see first qqplot below). Derivative values were transformed by raising the estimates to the 1/4th power. This produced a distribution of estimates that was more liky to match the assumptions of MLM (see second qqplot below). The forward selection MLM was fit to the transformed values.


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
0.0
0.5
1.0
1.5
-4
-2
0
2
4

¼ power


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
0.0
0.2
0.4
0.6
0.8
1.0
-4
-2
0
2
4

The final model selected by the stepwise algorithm included a threeway interaction between "Drive Condition", "Intercept Change" (due to call onset) and "Task Difficulty." There was also a two-way interaction between "Dive Condition" and "Time". The main effects for "Longitudinal Velocity" and "$Time^2$" were also significant. There was a problem in the model reaching proper convergence, as all of the correlations between random effects were not estimated, and the random effects all showed variances equal to zero. (Following table.)

```
Linear mixed model fit by REML
Formula: absLatVel ~ (1 + Time + I(Time^2) + IntChange + SlopeChange |
Sub) + LongVel + DriveCond + IntChange + Task + Time + I(Time^2) +
DriveCond:IntChange + DriveCond:Task + IntChange:Task + DriveCond:Time +
DriveCond:IntChange:Task
   Data: FullSubData
    AIC    BIC logLik deviance REMLdev
 -72588 -72313  36326   -72854  -72652
Random effects:
 Groups   Name          Variance   Std.Dev.   Corr
 Sub      (Intercept)   0.0000e+00 0.0000e+00
          Time          0.0000e+00 0.0000e+00   NaN
          I(Time^2)     0.0000e+00 0.0000e+00   NaN   NaN
          IntChangeTRUE 0.0000e+00 0.0000e+00   NaN   NaN   NaN
          SlopeChange   2.2738e-20 1.5079e-10   NaN   NaN   NaN   NaN
 Residual               9.6095e-03 9.8028e-02
Number of obs: 40306, groups: Sub, 1

Fixed effects:
                                          Estimate Std. Error t value
(Intercept)                              5.115e-01  9.652e-03   52.99
LongVel                                  1.745e-03  9.192e-05   18.99
DriveCondCongested                       9.729e-03  6.147e-03    1.58
IntChangeTRUE                            7.193e-03  3.898e-03    1.85
TaskSV                                  -1.436e-02  1.983e-03   -7.24
TaskCV                                   1.856e-03  2.007e-03    0.92
Time                                     1.955e-03  2.602e-04    7.51
I(Time^2)                               -3.201e-05  5.772e-06   -5.55
DriveCondCongested:IntChangeTRUE         1.164e-03  4.731e-03    0.25
DriveCondCongested:TaskSV                1.204e-02  2.755e-03    4.37
DriveCondCongested:TaskCV               -5.496e-03  2.781e-03   -1.98
IntChangeTRUE:TaskSV                    -2.093e-02  3.995e-03   -5.24
IntChangeTRUE:TaskCV                    -2.989e-02  4.052e-03   -7.38
DriveCondCongested:Time                  4.581e-04  1.235e-04    3.71
DriveCondCongested:IntChangeTRUE:TaskSV  1.014e-04  5.562e-03    0.02
DriveCondCongested:IntChangeTRUE:TaskCV -7.805e-03  5.629e-03   -1.39

Correlation of Fixed Effects:
              (Intr) LongVl DrvCnC InCTRUE TaskSV TaskCV Time   I(T^2)
DrCC:ICTRUE DCC:TS DCC:TC ICTRUE:TS ICTRUE:TC DrCC:T DCC:ICTRUE:TS
LongVel       -0.946
DrvCndCngst   -0.897  0.853
IntChngTRUE   -0.042 -0.032 -0.078
TaskSV        -0.011 -0.102  0.082  0.271
TaskCV        -0.039 -0.072  0.106  0.266   0.527
Time          -0.191 -0.082  0.030  0.290   0.008  0.006
I(Time^2)      0.161  0.061  0.052 -0.476  -0.006 -0.004 -0.940
DrCC:ICTRUE   -0.042  0.017  0.101 -0.638  -0.222 -0.219  0.128  0.001
DrvCndC:TSV    0.002  0.080 -0.161 -0.195  -0.720 -0.380 -0.007  0.005  0.300
```

```
DrvCndC:TCV    0.044  0.035 -0.198 -0.192  -0.378 -0.720 -0.003  0.003  0.296
0.511
InCTRUE:TSV    0.046  0.008 -0.077 -0.545  -0.492 -0.258 -0.001  0.001  0.449
0.354  0.186
InCTRUE:TCV    0.053  0.000 -0.083 -0.537  -0.257 -0.493  0.000  0.000  0.442
0.185  0.355  0.523
DrvCndCng:T    0.123  0.012 -0.402  0.332  -0.001 -0.001 -0.243  0.000 -0.535
0.001  0.000  0.000     0.000
DCC:ICTRUE:TS -0.027 -0.012  0.103  0.391   0.354  0.186  0.001 -0.001 -0.607
-0.493 -0.252 -0.718    -0.376     0.000
DCC:ICTRUE:TC -0.040  0.002  0.114  0.386   0.185  0.355  0.000  0.000 -0.600
-0.251 -0.493 -0.377    -0.720     0.000  0.510
```

A reduced data set was created by selecting every 3rd line in the original data (40306 rows, reduced data set had 13436 rows) to aid convergence. The final model was run on the reduced data set, and the significance of each of the key effects was compared in the Full Data set which did not converge properly, and the Reduced Data set which did converge properly.

Reduced Data Set --- Checking significance of key effects

Effect	AIC	BIC
Longitudinal Velocity	Large increase in AIC if this effect is removed; reduced data support this effect.	Same as AIC
Time^2	Large increase in AIC if this effect is removed; reduced data support this effect.	Same as AIC
Drive Condition x Time	Small increase in AIC if this effect is removed; reduced data support this effect.	Small decrease in BIC if this effect is removed; reduced data may not full support this effect.
Drive Condition x Intercept Change x Task	Small decrease in AIC when this effect is removed. Suggests this effect may not have strong support.	Reasonable decrease in BIC. Reduced data does not support this effect.

The analysis with the reduced data set did produce estimates of variances and covariances for the random effects, suggesting better convergence. Not all of the effects were fully supported in the reduced data. This could be due to a loss of power, but also could be due to changes in the standard errors due to non-zero variances of the random effects.

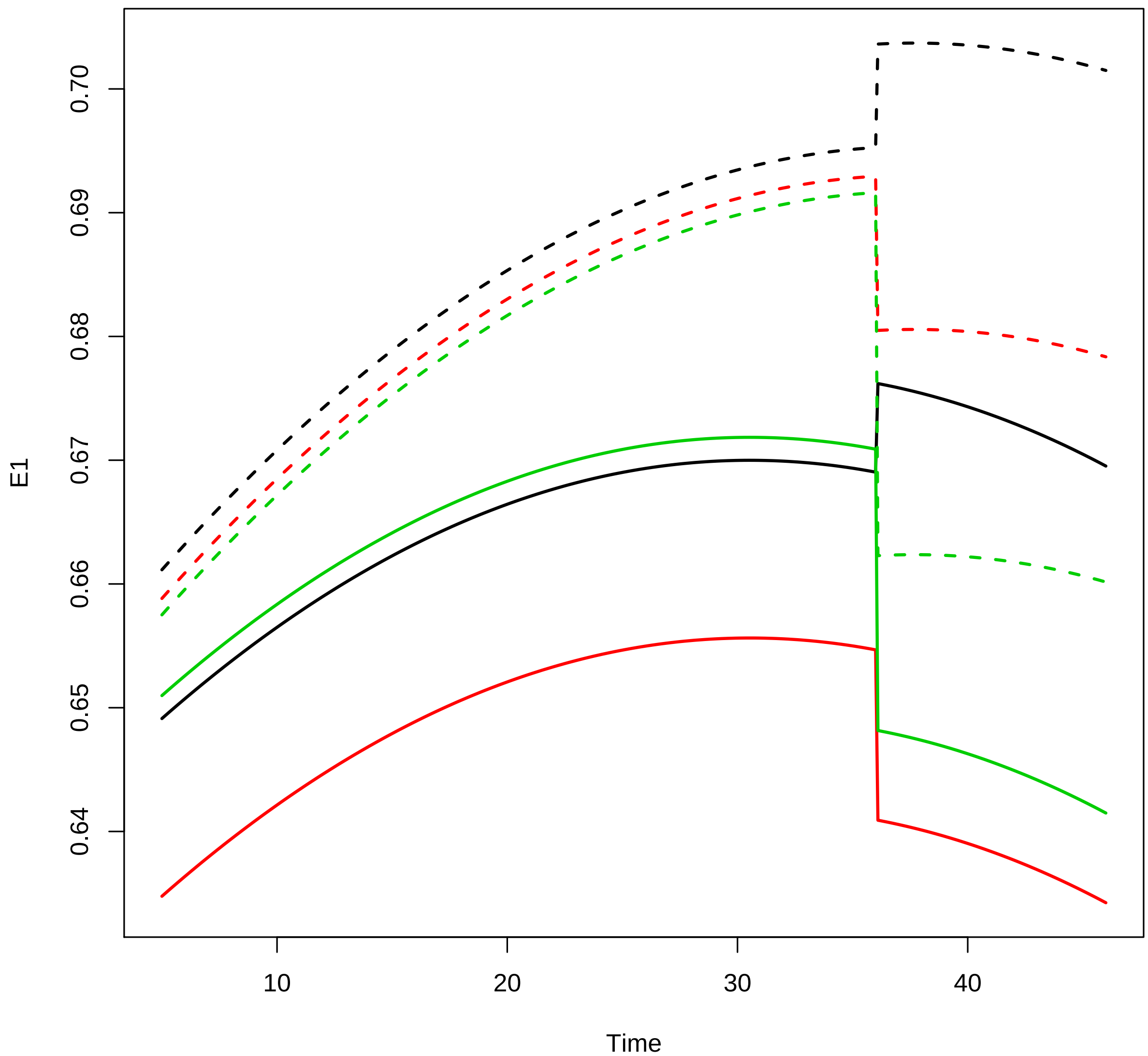
Removing the three-way interaction, which was not clearly supported in the Reduced Data, the Full Data set produced estimates of the variances and covariances of the random effects, but still did not converge properly. Based on AIC elements from the full model were removed until no further reduction in the AIC was observed. This yielded a model with a two way interaction between “Task Difficulty” and “Drive Condition”, a two way interaction between “Intercept Change” and “Task”, a two way interaction between “Dive Condition” and “Time” and a main effect for “Longitudinal Velocity”. There were still some odd values in the results, but these appear to come closer to reasonable convergence than the prior Full Data results with the three-way interaction.

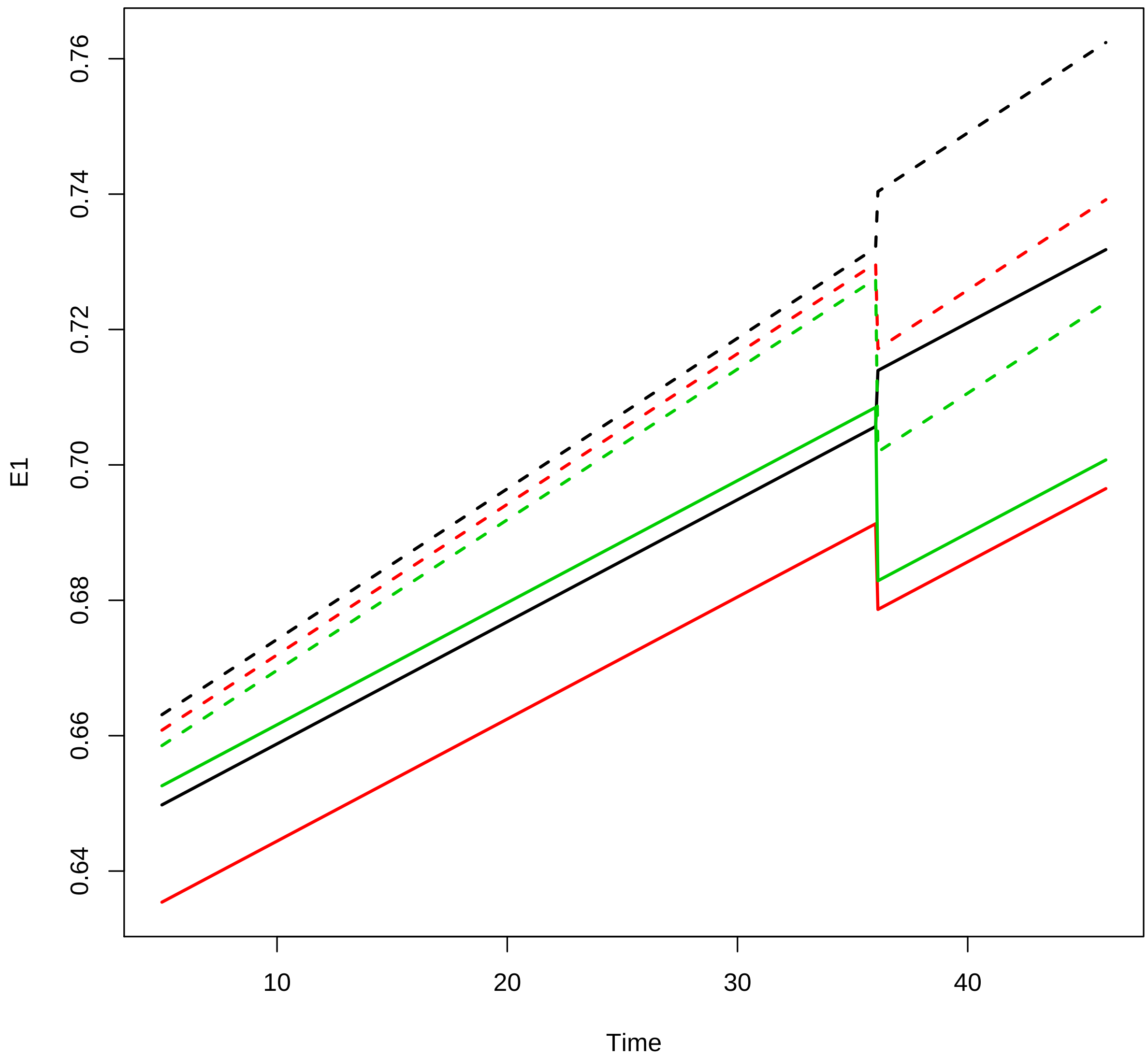
```
Linear mixed model fit by REML
Formula: absLatVel ~ LongVel + Task + Time + (1 + Time + I(Time^2) +
IntChange +      SlopeChange | Sub) + DriveCond + IntChange + DriveCond:Task
+      IntChange:Task + DriveCond:Time
   Data: FullSubData
    AIC    BIC logLik deviance REMLdev
 -72638 -72397  36347   -72842  -72694
Random effects:
 Groups   Name          Variance   Std.Dev.   Corr
 Sub      (Intercept)   1.0456e-18 1.0225e-09
          Time          2.7609e-18 1.6616e-09 -0.749
          I(Time^2)     7.9705e-10 2.8232e-05  0.000  0.000
          IntChangeTRUE 1.6590e-07 4.0731e-04  0.000  0.000  1.000
          SlopeChange   8.3223e-08 2.8848e-04  0.000  0.000  1.000  1.000
 Residual               9.6094e-03 9.8027e-02
Number of obs: 40306, groups: Sub, 1

Fixed effects:
                            Estimate Std. Error t value
(Intercept)                5.118e-01  9.572e-03   53.46
LongVel                    1.750e-03  9.187e-05   19.05
TaskSV                    -1.437e-02  1.854e-03   -7.75
TaskCV                     2.823e-03  1.876e-03    1.50
Time                       1.803e-03  2.291e-04    7.87
DriveCondCongested         1.122e-02  5.882e-03    1.91
IntChangeTRUE              8.092e-03  3.044e-03    2.66
TaskSV:DriveCondCongested  1.208e-02  2.397e-03    5.04
TaskCV:DriveCondCongested -7.394e-03  2.419e-03   -3.06
TaskSV:IntChangeTRUE      -2.093e-02  2.778e-03   -7.53
TaskCV:IntChangeTRUE      -3.389e-02  2.812e-03  -12.05
Time:DriveCondCongested    4.212e-04  8.264e-05    5.10

Correlation of Fixed Effects:
            (Intr) LongVl TaskSV TaskCV Time   DrvCnC ICTRUE TSV:DC TCV:DC
TSV:IC TCV:IC
LongVel     -0.954
TaskSV      -0.002 -0.105
TaskCV      -0.027 -0.077  0.525
Time        -0.161 -0.085  0.010  0.008
DrvCndCngst -0.915  0.887  0.053  0.075 -0.009
IntChngTRUE -0.079 -0.026  0.174  0.172  0.486 -0.019
TskSV:DrvCC -0.013  0.085 -0.671 -0.354 -0.007 -0.133 -0.002
TskCV:DrvCC  0.029  0.041 -0.353 -0.671 -0.004 -0.171 -0.002  0.511
TsSV:ICTRUE  0.039  0.000 -0.365 -0.186 -0.003 -0.007 -0.470  0.000  0.000
TsCV:ICTRUE  0.036  0.002 -0.186 -0.366 -0.004 -0.004 -0.465  0.000  0.000
0.509
Tm:DrvCndCn  0.078  0.037 -0.011 -0.009 -0.188 -0.325 -0.012  0.003  0.002
0.019  0.015
>
```

The results for the full model (with 3-way interaction) and reduced model (trying to correct for convergence issues) are respectively plotted in the top and bottom figures below. While the models differ substantially, the general pattern of results is similar (e.g., effect of the call late in the task, effects of different task difficulty). In both figures the black, red, and green lines correspond to the “Task Difficulty” levels of no-verbal, simple verbal, and complex verbal tasks respectively. The solid and dashed lines correspond to the monotonous and congested conditions respectively.





Notes for follow-up analysis

After the fact I noticed that the forward selection *may* have used REML rather than ML in the estimation process. If this is the case, the results would need to be rerun, as this might cause some small differences. Using ML would allow for a likelihood ratio test to be used, while the results from REML (last time I checked) do not allow for a likelihood ratio test to be performed.

Section 4a: Fractal Analysis

Summary of Analyses

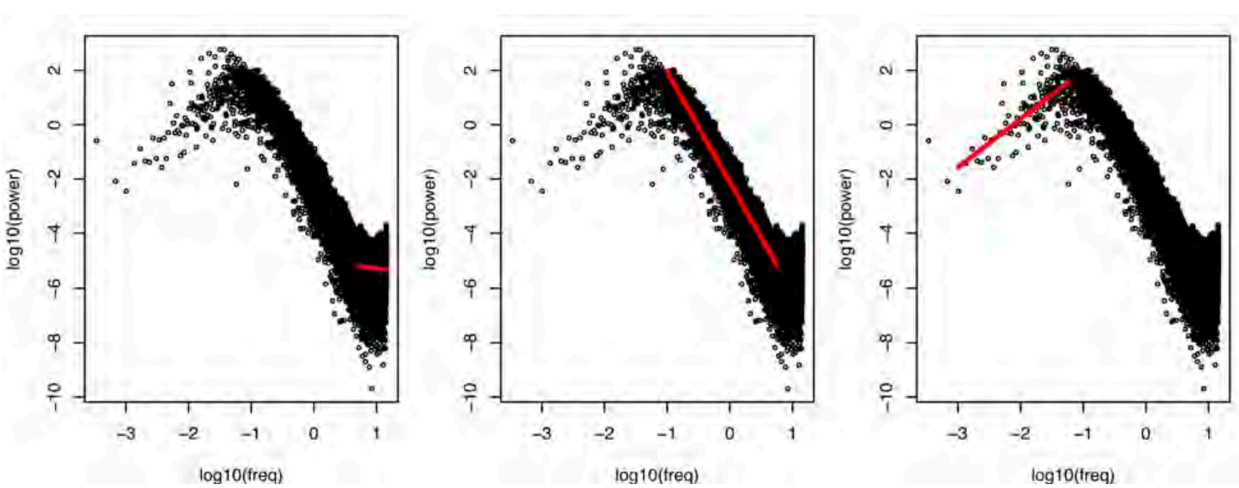
One way to gauge how smoothly a driver is or is not controlling a car is by considering whether the driver is making lots of corrections while driving or not. Plotting a measure such as lateral velocity over time, a driver making more corrections will produce a line with a more jagged profile than a driver making fewer, smooth changes while driving. One way to quantify the jaggedness of these lines is using tools developed for estimating fractal dimension. These tools come from the perspective that a jagged line in some sense falls between the 1-dimension nature of a straight line and the 2-dimensional nature of a plane; that is that a jagged line exists in a fractional dimension between a straight line and a plane.

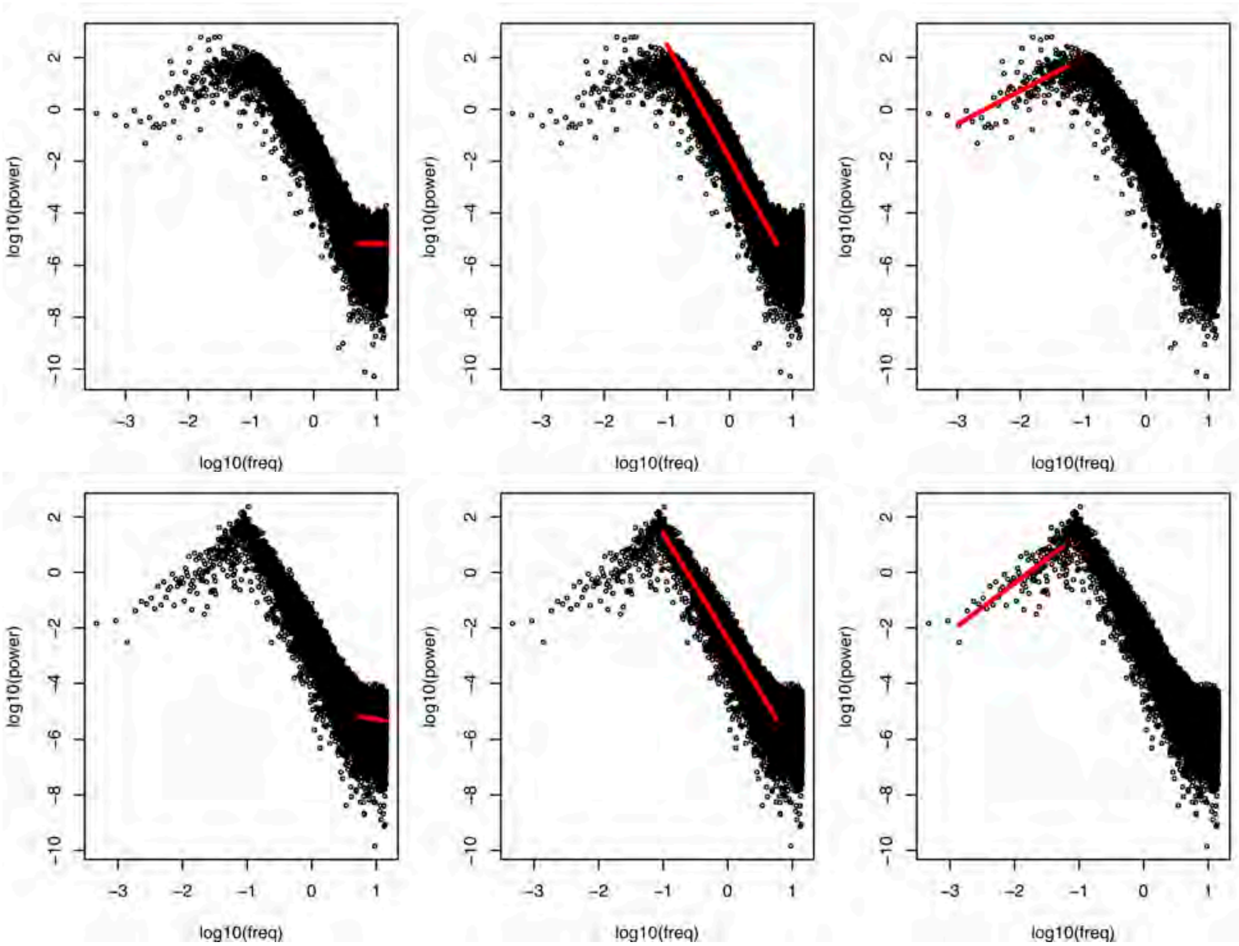
It has been shown that by calculating the spectral density of a time series, and producing a log-log plot of power to frequency, one can estimate a parameter related to the fractional dimension by estimating the slope(s) of the logged spectral density values. These resulting slopes can also be understood as relating to different forms of noise, which may differ from white noise in that they show properties such as longer term dependencies.

This analysis began by examining the log-log plot for the entire time series of several subjects to examine whether expected changes in slope in the log-log plot were apparent (more details to follow). Subsequently, individual time series were divided into smaller pieces, each of which was used to estimate fractional dimension parameters. These parameters were then correlated with lane deviance, a common measure of attention/performance.

Initial Log-Log Plots

The following figures are samples of the results observed for three subjects when the log-log plots were estimated using full time series. Each row of figures corresponds to a single subject, the three columns highlight three distinct regions which appeared to occur for most subjects. These regions correspond to variance in lateral velocity occurring over differing time scales in the time series.





Log10(freq) less about 0.75 tended to produce slopes approximately equal to zero (first column of figures). This range of frequencies corresponds to oscillations occurring over approximately 0.2 seconds. The slope near zero suggests these values behave as one would expect for white nose (normally distributed, independent observations), and are likely mostly attributable to non-human factors such as electronic hiss.

Log10(freq) from about 0.75 to -1 tended to produce steep negative slopes (middle column of results). The negative slopes in the examples provided range from -3.8 to -4.4. The steepness of these slopes suggest persistent Brownian motion --- close to the realm of processes such as random walks. This suggests high correlations between subsequent observations. The frequencies in this range correspond to oscillations occurring over 0.2 to 10 seconds. This range seems to closely correspond to the behavior expected for controlled changed in lateral velocity of a car, and was approximately the range and slope expected prior to analysis.

Log10(freq) greater than about -1 tended to produce a positive slope (last column). Slopes for the three examples provided ranged from 1.2 to 1.8. The range of frequencies corresponds to osccillations occuring over the course of 10 or more seconds, and upwards of several minutes. This range is likely to correspond to large scale features of the data such as curves in the road.

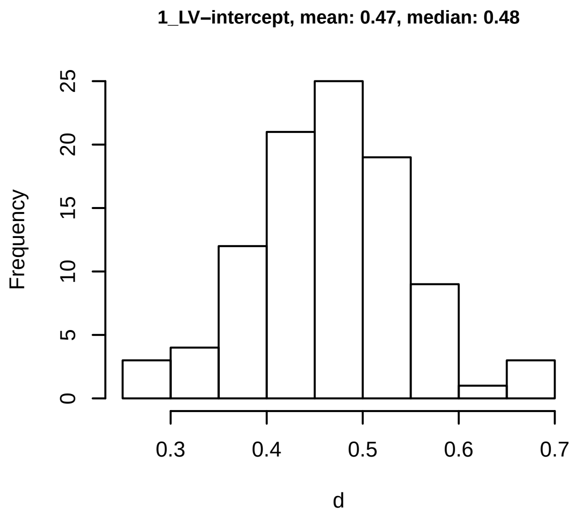
Comparison of Fractional Dimension Information to Lane Variance

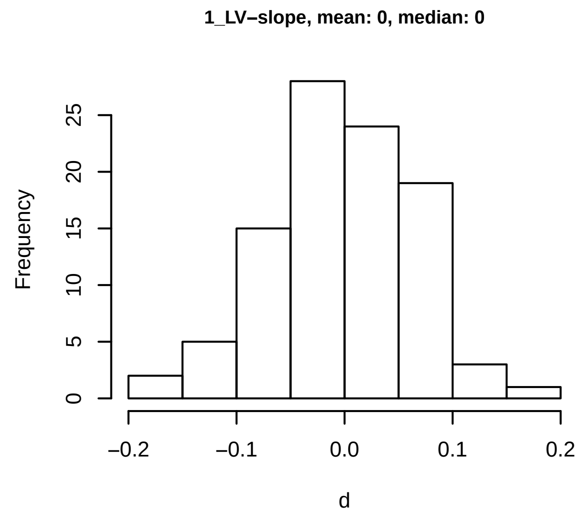
Subsequent analysis focused on the middle range of values (fractions of a second to a few seconds), as this range seemed most likely to correspond to ongoing driver performance. The lower range (<0.2 seconds/oscillation) and upper range (>10 seconds/oscillation) were not examined, as these were more likely to relate to characteristics of the equipment and the simulation conditions respectively. The slope of the middle range was estimated using Log10 values from -0.3 to 0.3 correponding to oscillations in the range of 0.5 to 2 seconds. Other ranges were examined, but those are not presented hear as pronoced differences in the results were not observed.

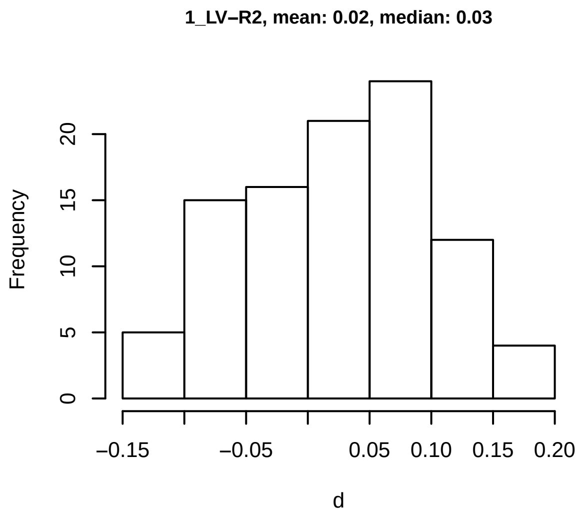
For the middle range over which the slope was estimates, the intercept, slope and R^2 of the linear regression (i.e., red lines in figures) was recorded. Based on the spectral density, the percentage of power contained in the range examined was also calculated. The intercept of the linear regression can be understood as relating to the amplitude of the lateral velocity movements for oscillations of about one second. The slope of the linear regression can be understood to capure how jagged the changes in lateral velocity are. The interpretation of the R^2 for the linear regression give some impression as to how well the linear regression is capturing the information in the log-log plot over the range examined, and may reveal minor variations in whether the fractional dimension is consistent in the range examined. The percentage of power calculation gives an impression as to the proportion of the total variation in lateral velocity occurs in the range examined. Because power was very skewed, log(power) was also considered.

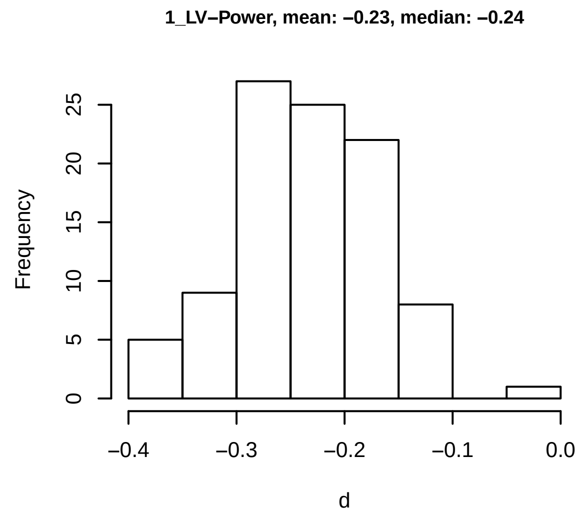
Individual time series were divided into smaller intervals, so as to examine the change in estimated parameters over time. While a large range of intervals was examined, in the end intervals of 0.5, 1, 2, and 5 minutes were examined. Lane variance estimates (relative to the center of the lane to which participants were closest) were calculated over the same interval of time. Correlations between lane variance and the five measures listed above were calculated for each individual separately. Log of lane variance was also considered, due to how skewed lane variance was within some windows of estimation.

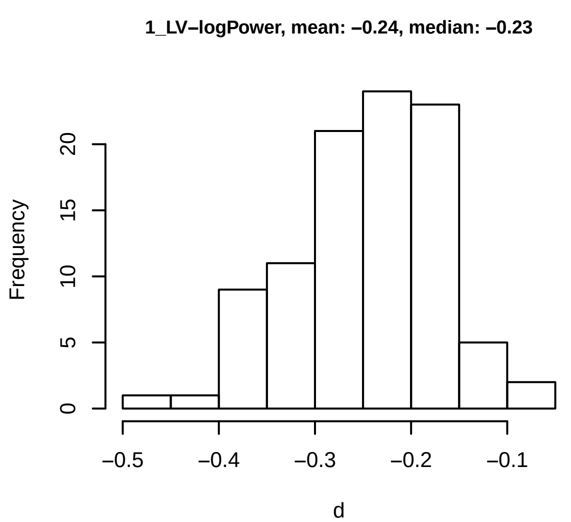
The figure below is one of 10, but gives an example of the results examined. Figure [1,1] (row 1, column 1) is the correlations for each of the individual participants between lane variance and the linear regression intercept from the log(freq)-log(spectral power) plots above; these results are specific to the log(freq) range of -.3 to .3, and time series windows of 0.5 minutes. Figure [1,2] are the correlations between lane variance and slope. Figure [2,1] are the correlations between lane variance and R^2. Figure [2,2] are the correlations between lane variance and percentage power. Figure [3,1] are the correlations between lane variance and log(power).

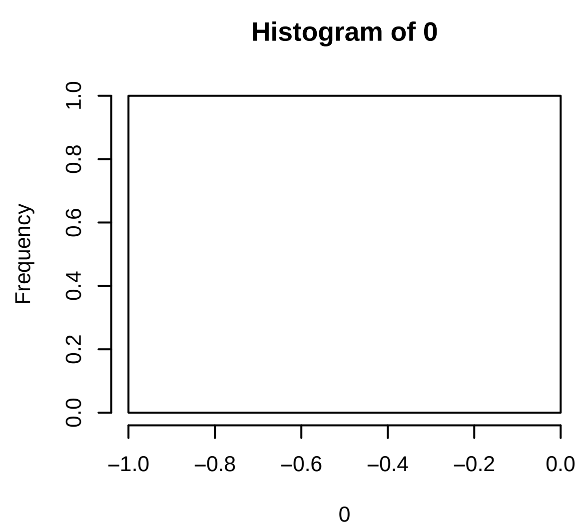

1_LV–intercept, mean: 0.47, median: 0.48
Frequency
d
0 5 10 15 20 25
0.3 0.4 0.5 0.6 0.7


1_LV–slope, mean: 0, median: 0
Frequency
d
0 5 10 15 20 25
–0.2 –0.1 0.0 0.1 0.2


1_LV–R2, mean: 0.02, median: 0.03
Frequency
d
0 5 10 15 20
–0.15 –0.05 0.05 0.10 0.15 0.20


1_LV–Power, mean: –0.23, median: –0.24
Frequency
d
0 5 10 15 20 25
–0.4 –0.3 –0.2 –0.1 0.0


1_LV–logPower, mean: –0.24, median: –0.23
Frequency
d
0 5 10 15 20
–0.5 –0.4 –0.3 –0.2 –0.1


Histogram of 0
Frequency
0
0.0 0.2 0.4 0.6 0.8 1.0
–1.0 –0.8 –0.6 –0.4 –0.2 0.0

Over all conditions examined, the slope and R^2 estimates generally produced low correlations with lane variance. Extensive examination of nonlinear relationships was not examined, as monotonically increasing/decreasing relationship was expected between these variables; larger correlations would be expected for any such relationship.

The intercept parameter and power estimates seemed to produce the most promise as parameters that would be related to momentary changes in driver attention. The correlations between power and lane variance were much more variable than the intercept term; this was likely due to the fact that the power estimate will depend on other frequencies in the data. As the presence of greater/lesser contributions of other frequencies would increase/decrease power, it was decided that the interpretation of the power estimates might be less succinct that the interpretation of the estimated intercept parameters.

MLM in the following section focused on the modeling of the intercept.

Side note: It took a little while to see this, but I can demonstrate that the intercept is mathematically related to the average amplitude of components in the frequency range examined, as spectral power is related to the $(\text{amplitude-of-the-frequency})^2$.

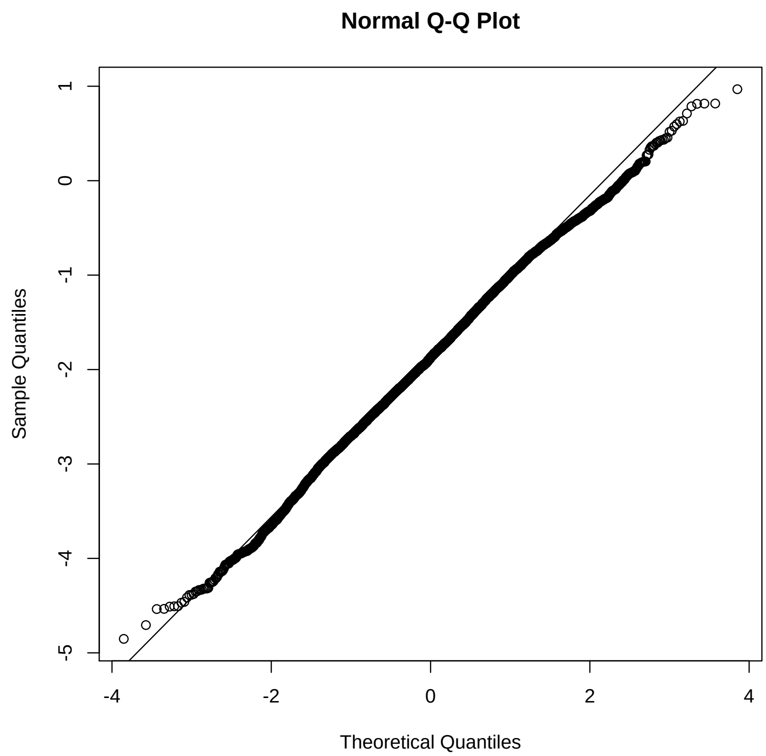
Section 4b: Multilevel models of fractal estimates

Summary of Analyses

Based on the prior analyses, the intercept parameter was estimates over non-overlapping intervals of 0.5, 1, and 2 minutes. Intervals were estimated for each individual over the course over the time series; the first 5 minutes of the time series were removed. A multilevel model (MLM) was fit to the data to allow individuals to differ in their parameter estimates; for example the effect of time was allowed to differ between individuals.

Data Preparation & Analysis Notes

Double checked the distribution of the dependent variable. Based on a quantile-quantile normal plot, the dependent variable appears to be very close to a normal distribution.



The model was run allowing for the following predictors:

- -Drive Condition: congested/monotonous
- -Task Difficulty: No Verbal, Simple Verbal, Complex Verbal
- -Time
- -Time^2: Allowed the effect of time to increase/decrease over time
- -Longitudinal Velocity: Lateral velocity is related to longitudinal velocity, so this can be considered a covariate
- -Intercept Change due to onset of telephone call

-Slope Change due to onset of telephone call

The model was run allowing for the following interactions:

- -Drive Condition x Task Difficulty
- -Drive Condition x Time
- -Drive Condition x Time^2
- -Drive Condition x Intercept Change
- -Drive Condition x Slope Change
- -Task Condition x Intercept Change
- -Task Condition x Slope Change
- -Drive Condition x Intercept Change x Task Difficulty
- -Drive Condition x Slope Change x Task Difficulty

Random Effects were allowed for the following predictors:

- -Intercept
- -Time
- -Time^2
- -Intercept Change
- -Slope Change

A forward selection algorithm was used to select the ideal set of predictors; this seemed reasonable as initial models suggested that the significant amount of data tends to produce either very significant or clearly non-significant results. The addition of the next variable was based on likelihood improvement. The algorithm used the likelihood ratio test to decide when to stop. For the final model, AIC was used to confirm that there was no evidence suggesting the need for additional predictors. The initial model consisted of an intercept and all of the random effects.

Results

The following results are based on estimates made over the course of 0.5 minutes.

The final model suggested significant effects for:

- -Drive Condition x Time
- -Drive Condition x Time^2
- -Task Condition x Intercept Change

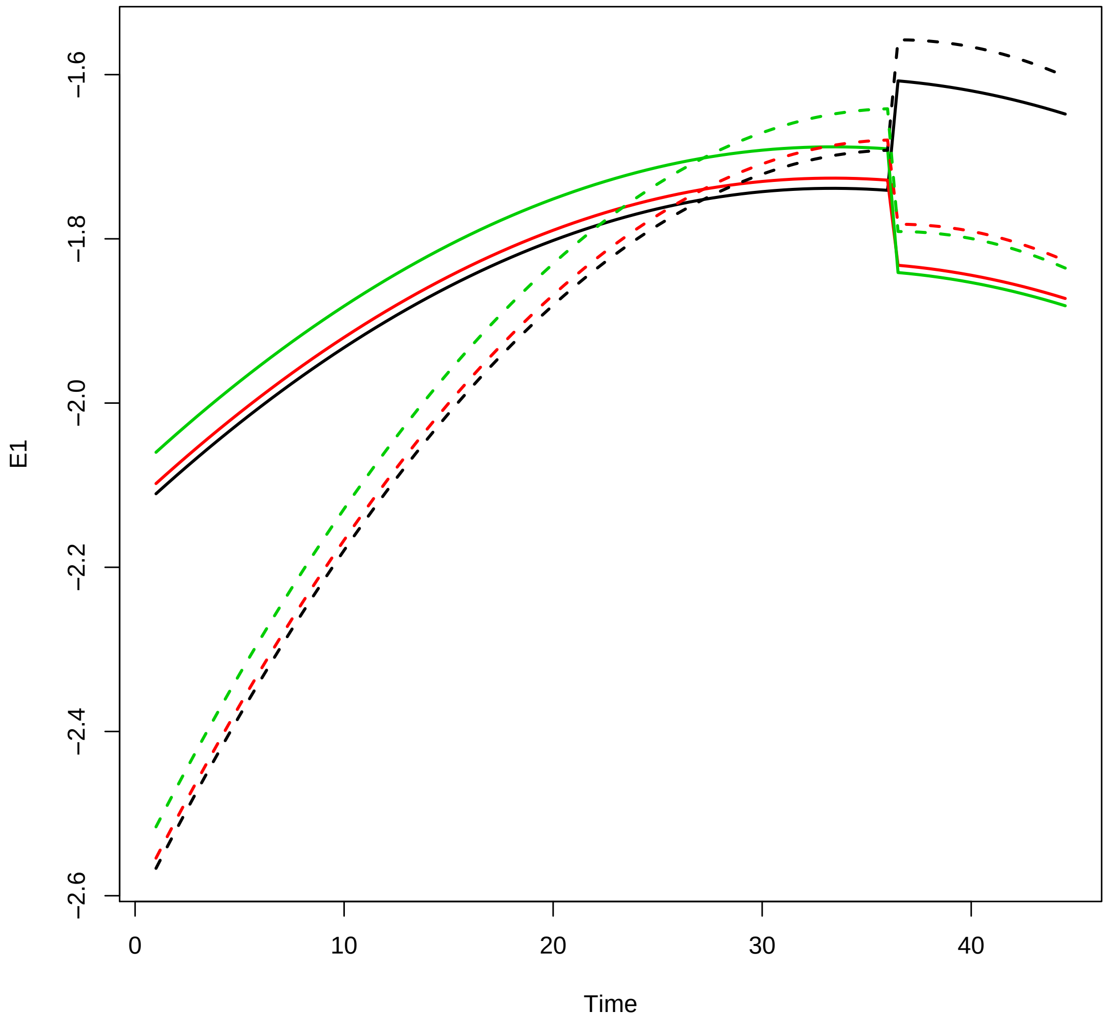
All lower order effects were included in the model. Longitudinal Velocity was included in the model as a covariate, and was also shown to be a significant predictor to the dependent variable that is based on Lateral Velocity. The random effects all suggested substantial individual differences in the effects of Intercept, Time, Time^2, Intercept Change and Slope Change. The table below summarizes the results.

```
   Data: FullSubData
   AIC   BIC logLik deviance REMLdev
 13345 13542  -6644    13193   13289
Random effects:
 Groups   Name          Variance   Std.Dev.   Corr
 Sub      (Intercept)   1.4680e-01 0.38314364
          Time          1.3650e-03 0.03694618 -0.483
          I(Time^2)     7.4854e-07 0.00086518  0.457 -0.960
          IntChangeTRUE 9.4646e-02 0.30764650  0.122 -0.012 -0.169
          SlopeChange   8.3222e-04 0.02884829 -0.409  0.766 -0.842 -0.032
 Residual               2.5236e-01 0.50235468
Number of obs: 8536, groups: Sub, 97

Fixed effects:
                               Estimate Std. Error t value
(Intercept)                  -3.1830269  0.1819047 -17.498
LongVel                       0.0142830  0.0016196   8.819
DriveCondCongested           -0.4828460  0.1353486  -3.567
Time                          0.0236876  0.0048454   4.889
I(Time^2)                    -0.0003549  0.0001035  -3.429
TaskSV                        0.0124217  0.0858253   0.145
TaskCV                        0.0505198  0.0878494   0.575
IntChangeTRUE                 0.1343419  0.0534588   2.513
DriveCondCongested:Time       0.0269913  0.0067303   4.010
TaskSV:IntChangeTRUE         -0.2370191  0.0548646  -4.320
TaskCV:IntChangeTRUE         -0.2839938  0.0561530  -5.058
DriveCondCongested:I(Time^2) -0.0003395  0.0001328  -2.556
```

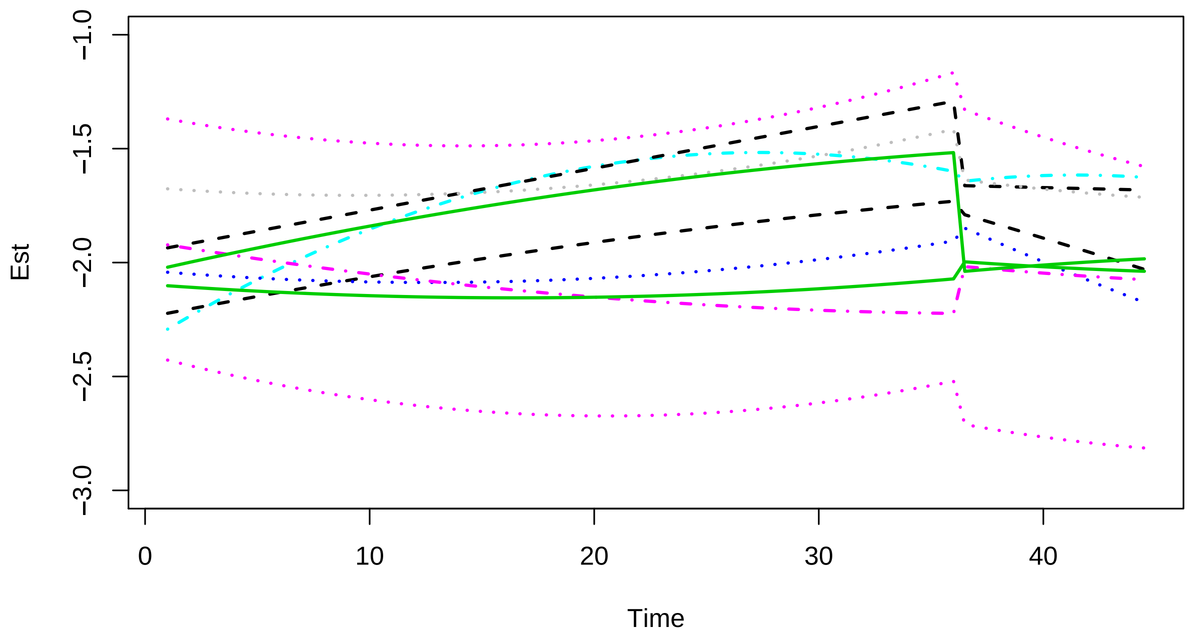
The figure below summarizes the fixed effects of this model. The solid lines correspond to the monotonous condition, while the dashed lines correspond to the congested condition. The congested condition started with lower estimates that the monotonous condition, but showed a larger long-term effect of time. The effect of time was non linear, such that the increase in the dependent variable was steepest early on in the drive.

The black, red and green lines correspond to the “No Verbal Task”, “Simple Verbal Task” and “Complex Verbal Task”. An intercept increase was observed for the “No Verbal Task”, while decreases in the dependent variable were observed for the “Simple Verbal Task” and “Complex Verbal Task”.

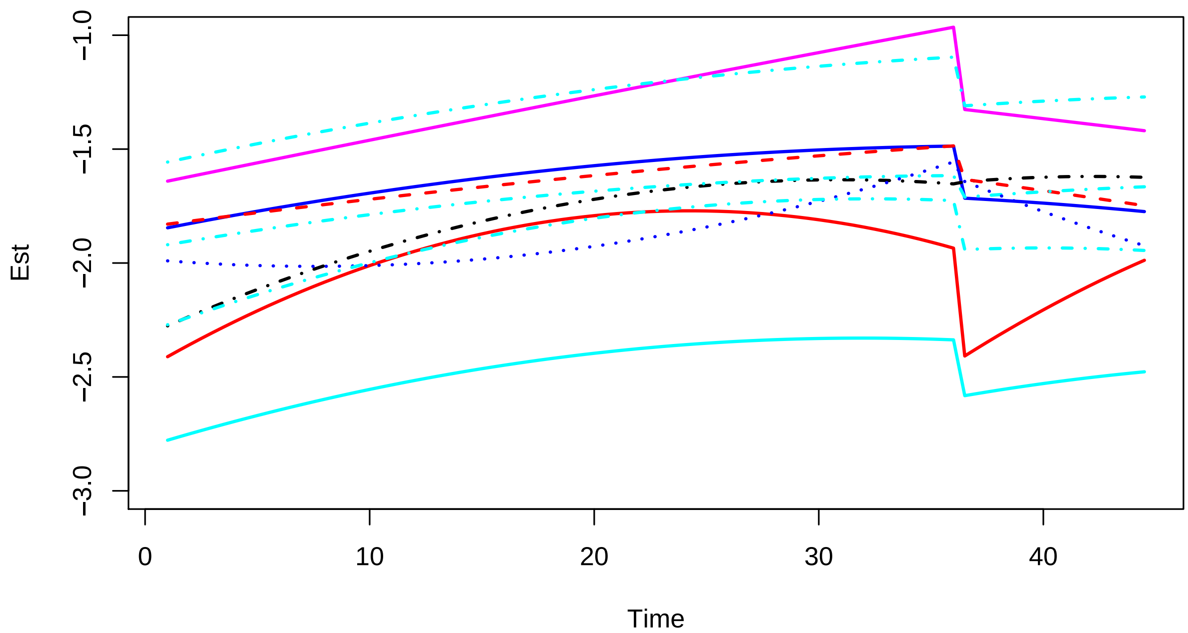


The following two plots are examples of the individual estimates of trajectories, taking into account the random effects estimated for a sample of 10 individuals. Both plots are for the complex verbal task, and represent the congested and monotonous driving conditions. Not explored here are whether the differences between individuals (i.e., the random effects) are systematically related to any inter-individual measurements, as no inter-individual variables were provided. Colors and line types are meaningless; they were selected randomly to differentiate individual trajectories.

Individual Plots, Congested, Complex Verbal



Individual Plots, Monotonous, Complex Verbal



Section 5a: Relating reaction time task to momentary derivative estimates

Summary of Analyses

In the prior MLM analyses using momentary derivative estimates, the resulting model behaved in ways that were expected for variation in attention over the course of the drive; there were increases in variability of momentary derivative estimate associated with time and relationships with other events such as the onset of a telephone call during the last ~10 minutes of the drive.

The present analysis aims to link the results from an ongoing attention task with the momentary derivative estimates. Such a link would demonstrate a closer connection between momentary derivatives and ongoing attention processes.

Data Preparation & Analysis Notes

Aligning data: PDT task was started at 19,500 feet for condition 1 (monotonous condition) and 11,750 feet for condition 2 (congested condition); PDT and driving task were stopped at the same time. Time was dilated/contracted so that the final time points matched and that to take into account the beginning of each time series. In general the amount of time dilation/contraction appears to have been small, so including this correction (or not) shouldn't make a very big difference on the results.

Only a very small percentage of wrong responses were recorded (~2%). Consequently, using right/wrong responses as a predictor or outcome is unlikely to be productive. Focus was consequently placed on analyzing the reaction times provided when subjects were given a "go" task.

The PDT and driving task were recorded at different sampling rates. The time on the driving task was matched to times on the PDT task for all occasions were a reaction time was recorded. A maximum allowable deviation of 0.1 minutes was used; a slightly wider allowable interval was used as there were some gaps in the time scale for the driving task.

Estimates of the absolute value of the first derivative of lateral velocity were made approximately every 2 seconds of the time series.

Primary analysis: A MLM was estimated with the derivative estimates as the dependent variable. Models with and without random effects for reaction time were considered; a random intercept model was used in all cases.

Other attempts: Monotonous and congested conditions were considered together. Several follow-up plots/analyses were conducted to check if any differences occurred if the tasks were separated. Considered both raw estimates of the derivative made every 2 seconds, as well as loess smooth derivative estimates with smoothing occurring over 1-5 minutes; this was done to both reduce error and to get a better estimate of the local performance of and individual.

Results

No models showed any clear indication of a significant result for the fixed effect of reaction time. P-values were generally well beyond marginal. The few things that came close to marginally significant were very sensitive to choices made in the analysis such as the smoothing parameter in the loess-smoothed results; this sensitivity to the analysis suggests that the evidence for a relation between the ongoing estimates of reaction time and the momentary estimates of lateral velocity derivatives is not particularly strong. In many cases the marginal results also appeared to be tied to the skewed nature of the dependent variable --- a violation of the MLM assumptions.

The reaction time measure was relatively noisy and infrequent. As the derivative measure was also relatively noisy, the power for finding a relation between these two measures may have been very low. Smoothed versions of the derivative variable did not appear to ameliorate the problem.

Other Analyses

It as briefly considered whether the slopes of RT and the Derivative estimates were related. Each variable was predicted using time, and the slopes recorded. There was no clear evidence that the linear slopes were correlated, which was not surprising as the correlation was based on a small number of subjects (n=18).

Section 5b: Relating reaction time task to fractal estimates

Summary of Analyses

Like the momentary derivative estimates, it was considered whether the fractal estimates could be predicted over the course of the driving task using the reaction times collected on the PDT task. The goal was once again to try to identify an ongoing relationship between these variables.

Data Preparation & Analysis Notes

All data preparation and analysis notes from Section 5a apply in this section. Similar MLMs were considered. While the dependent variable in the previous analysis was relatively skewed, the dependent variable in this analysis was close to normally distributed and therefore the assumptions of MLM should be better met in these analyses.

Results

Like the previous results, no clear evidence linking the reaction time measure with the dependent variable was observed. Smoothed fractal estimates appeared to produce some marginal results, but these were very sensitive to the smoothing parameter suggesting a lack of strong evidence for a link between the measures. Steps such as controlling for condition (congested versus monotonous) did not appear to provide an advantage for clarifying a relationship between RT and the dependent variable.

Additional work: Testing vigilance detection methods that are compatible with walking

The original proposal was designed to examine driving performance. However, there is also interest in detecting soldier readiness and vigilance for soldiers on foot. Loss of vigilance during sustained operations by soldiers on foot represents a grave potential threat to safety. What is needed is the ability of a commander at a local level to have a cognitive readiness profile of their soldiers during the course of sustained activities. This measurement should be able to be accomplished remotely, it should not interfere with other tasks, and it should not disrupt the security of the unit through increased light or sound.

The work described in the following section explores using a peripheral detection task (PDT) device to measure sustained attention. (NOTE: This PDT is NOT the version included in the original behavioral work.) The device used here presents a light to a participant that they must respond to by either clicking a small button or withholding a response. This device can also use vibrotactile stimulation instead of a light. In the present work, we used a driving task to bridge results and methods to present work. Our goal is to see if the device can predict losses in attention and then move to examine how well the device works for walking participants, instead of drivers.

Summary of results and conclusions:
The device proved effective at measuring declines in vigilance over time. However, these declines were only detectable using sophisticated hierarchical linear models. This implies that further work is needed to develop the best set of algorithms to analyze these type of data.

Promise for future work:
The PDT may be an excellent way to measure individual readiness of a soldier over time, and to develop a profile of unit cognitive readiness prior to an operation, after soldiers have been deployed in the field for a length of time. A commander should be able to query their soldiers and get an update within a brief period of time, without compromising unit security. Or, a command can be provided with an automatic readiness profile at pre-determined intervals to would provide them with an understanding of how readiness is changing. Additional work can help determine how to best realize this goal by developing the device under field conditions and building and testing the best measurements and data analysis techniques.

Section 1. Method

Participants (Drivers)

For this study, 20 students from the University of Kansas were recruited via flyer and advertisement, 9 males an 11 females (M = 22.7 years, SD = 2.6). Drivers were compensated $20 for their participation. All drivers reported having at least 3.5 years of driving experience. Almost all drivers reported English as their primary or native language; however failure to meet this requirement did not disqualify participants from the study. Drivers were randomly assigned to one of four conditions; monotonous drive with no verbal task, monotonous drive with verbal task, congested drive with no verbal task, or congested drive with verbal task.

Materials and Apparatus

Driving simulator

The driving scenario was designed using STISIM Drive (Systems Technology Inc. Hawthorne, CA) simulator software (Version 2.08.02). A fixed-base cockpit with force-feedback steering was used. The vehicle was set on automatic transmission. Drivers viewed the simulated roadway on a single 17-in LCD display. A Fresnel lens was placed between the LCD display and driver.

Scenarios

The roadway was a four-lane highway separated by a median. Both scenarios had a lane width of 3.66m, and visibility was set at 457.2m in clear conditions. The monotonous drive was generally flat with very few curves or hills, and no traffic in either direction. The monotonous scenario was designed to be under-stimulating to mimic the driving conditions that most motorists encounter while driving on a rural highway. Such conditions have been likened to a vigilance task (Papadelis et al., 2007; Thiffault & Bergeron, 2003). The monotonous drive was 350,000 feet in length. Drivers were instructed to drive between 45 and 50 mph, resulting in a consistent approximate drive time of 90 minutes. The congested drive was also generally flat, but had several curves, lots of buildings and street signs, and was congested with traffic in both directions. This congestion would prevent the driver from going a consistent speed limit. The congested drive was 211,000 feet in length. The consistent congestion plus a speed limit of 40mph resulted in an average speed throughout the drive of 25 to 30 mph, resulting in a consistent approximate drive of 90 minutes.

Peripheral Detection Task

The peripheral detection task was administered via the Detection Response Task headset (PDR Inc., Salt Lake City, UT) and accompanying software (version 3.0.23). A small LED light fixed on the end of a flexible arm is mounted on the participant's head, while a small response button is fixed to the participant's thumb or forefinger. The light is either red or green in color, controlled by the software. Reaction time is measured in microseconds.

Procedure

After completion of consent and a demographic data sheet that queried cell phone usage behavior, the experimenter asked drivers to turn their cell phones off and leave them in a room separated from the driving simulator to minimize potential distractions. Next, drivers were equipped with the peripheral detection headset with the light placed 6-8" from the head in the periphery of the left eye and a hands-free kit with the headphone over their right ear and the microphone extended over their mouth. Then drivers engaged in a practice drive that lasted approximately five minutes. This drive allowed drivers to familiarize themselves with the simulator and the handling of the steering controls. At the end of the practice drive, participants were given an example of the peripheral detection task (PDT), lasting approximately two minutes. Participants were instructed to react as quickly as possible to green lights and ignore the red lights. This practice allowed drivers to familiarize themselves with the detection task and ensure button placement was optimal for responses. This practice also allowed researchers to ensure that instructions for participating in the peripheral detection task were understood. Finally, drivers were able to practice the verbal task, and instructed that they may or may not be presented with this task during their drive. A second computer, responsible for the interactive verbal task would be activated remotely after the drive began. The experimenter would then move to the observation room, where the driver could be observed for the duration of the study to ensure that the driver was actively involved with the task. The experimenter did not converse with the driver during the experiment. After completing the drive, drivers completed an electronic version of the NASA-Task Load Index (Hart & Staveland, 1988). The entire experiment lasted approximately 110 minutes and was conducted in a darkened room with no ambient light source.

Performance Measures

Reaction Time (RT)

This measure refers to the time taken for the driver to respond via button press to the green light on the peripheral detection task headset. This measure was taken approximately 300 times per participant during the drive. This data was analyzed in two ways. An ANOVA compared average reaction time per participant across the four conditions (Howell, 2002). Additionally, this data was analyzed in an exploratory manner using multilevel modeling, nesting time at which RT was measured within participants.

Section 2. Results

ANOVA Analysis

Prior to the hierarchical linear modeling analysis, we used a 2x2 between-groups analysis of variance (ANOVA) as a comparative approach to answering our hypotheses. Reaction times to correct responses to the PDT were averaged across each participant. Means and standard deviations are summarized in Table 1. There was no significant interaction of the type of drive (1=congested vs.0= monotonous) and presence of the verbal task (1=present vs. 0=not-present), $F(1, 16) = 0.019$, $MSE = 15e^9, p = 0.893$. An investigation of the main effects also yielded no significant results. There was no significant main effect of type of drive, $F(1, 16) = 0.009$, $MSE = 13e^7, p = 0.926$. Likewise, there was no significant main effect of the presence of the verbal task, $F(1, 16) = 0.084$, $MSE = 12e^8, p = 0.775$. This analysis reflects the typical statistical approach that would have been conducted with this dataset.

HLM Analysis

In lieu of no statistically significant effects found using analysis of variance, hierarchical linear modeling was used to analyze the multilevel model specified by our hypotheses. We began by determining the total variance residing at the participant level. An intraclass correlation of 0.538 indicated that approximately 53% of the total variance resides between participants; as such, an HLM analysis of this data was certainly appropriate.

Baseline Model

It was important for us to construct a null model to serve as a basis of comparison for further, more complicated modeling approaches. Results from this model, which predicted reaction times (measured in microseconds) as a function of time progressed in the simulator (measured in seconds), are summarized in Table 2. There was no significant fixed effect of time on reaction times to the PDT, $\beta = 3.15$, $t = 1.425$, $p = 0.15$. However, as we were interested in the introduction of additional factors into subsequent models, we were not deterred by the non-significance of the fixed effect of time.

Model Building Procedure

Model construction progressed via the sequential addition of the type of drive and verbal condition parameters. In each successive model (wherein both verbal condition and type of drive were included), there was no significant fixed effect of the type of drive, the verbal condition, or time. This was also the case for all two-way and three-way interactions. Nested model deviance tests yielded no results; the successive models did not improve model fit over that of the null model. However, the exclusion of the verbal condition as a parameter did provide statistical significance in one case; we based our final analysis on this model (described below).

Time by Drive Type Model

A significant fixed effect of time and a marginally significant fixed effect of the time by drive type interaction were found during the model building process ($\beta = 6.658$, $t = 2.308$, $p = 0.02$; $\beta = -8.471$, $t = -1.89$, $p = 0.058$, respectively). There was no significant main effect of drive type, $\beta = 14877.22$, $t = 0.28$, $p = 0.779$. The results from this model are summarized in Table 3. However, this model did not provide statistically significantly different model fit from the null model, χ^2 (6) = 3.49, $p = 0.17$. At this point of the analysis, it was brought to our attention that a heteroscedastic error structure, wherein the residual variance for each successive time point may differ, would be more appropriate given that reaction times are more varied at the beginning of the drive.

Heteroscedastic Final Model

The results from fitting the heteroscedastic model are summarized in Table 4. This model elicited significant fixed effects of both time and the time by drive type interaction (β = 9.0, t = 17.05, p = 0.00; β = -9.5, t = -2.20, p = 0.02, respectively). There was no significant main effect of drive type, β =17254.4, t = 0.34, p = 0.7362. Contrary to the previous models, this model did provide significantly different model fit compared to the null model, χ^2 (25) = 57.38, $p < 0.001$.

Section 3. Implications

ANOVA Results

Without prior experience using hierarchical linear modeling, our analysis would likely have been conducted only using the 2x2 ANOVA explicated above. This analysis yielded no significant results regarding the two condition sets. As such, our exploration of using the PDT as a measure of vigilance may have been unnecessarily ended, as during the course of a pilot study we often interchange attentional measures dependent upon early results. Through this analysis, we found no effect of either the type of drive or the presence of a verbal task, which suggests that these two characteristics of a drive do not influence reaction times to a peripheral attention task, and therefore may not significantly increase the risk of being involved in a crash. Given knowledge of previous literature we would typically attribute the lack of a sufficient effect to a power issue.

HLM Results

Unlike the ANOVA analysis, the final hierarchical linear model yielded a significant fixed effect of time, as well as a significant fixed effect of the time by drive type interaction. The positive coefficient of time suggests that as participants continued throughout the drive, their reaction times on the PDT increased for each second increase in time. This further suggests that during an extended drive, attentional deficits increase as a linear function of the length of the drive, replicating Strayer and Drew (2003). Furthermore, the negative coefficient of the time by drive type interaction (where 1=congested drive) suggests that although there is a decrease in performance over time for all drivers (provided by the main effect of time), drivers on congested roadways demonstrate a slower rate of performance decrement than drivers on monotonous roadways.

Comparison of Hierarchical and ANOVA analyses

The ANOVA analysis did not yield significant results, whereas the HLM, after an extensive model-building process, did present substantive effects. Without using HLM, we likely would have overlooked these effects and simply attributed their lack of presence to a lack of sufficient power. However, with the more thorough analysis, we were able to obtain findings that will further direct this line of research. This is especially exciting as the effects were found solely within the pilot study, perhaps allowing for an increase in experimental complexity as we continue to develop the simulator methodology.